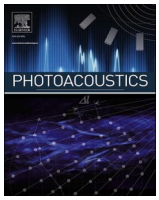
Photoacoustics 38 (2024) 100623

Contents lists available at ScienceDirect

Photoacoustics

journal homepage: www.elsevier.com/locate/pacs



Score-based generative model-assisted information compensation for high-quality limited-view reconstruction in photoacoustic tomography

Kangjun Guo [1], Zhiyuan Zheng [1], Wenhua Zhong [1], Zilong Li [1], Guijun Wang, Jiahong Li, Yubin Cao, Yiguang Wang, Jiabin Lin, Qiegen Liu, Xianlin Song [*]

School of Information Engineering, Nanchang University, Nanchang 330031, China

ARTICLE INFO

Keywords:
Photoacoustic tomography
Generative model
Limited view
Information compensation

ABSTRACT

Photoacoustic tomography (PAT) regularly operates in limited-view cases owing to data acquisition limitations. The results using traditional methods in limited-view PAT exhibit distortions and numerous artifacts. Here, a novel limited-view PAT reconstruction strategy that combines model-based iteration with score-based generative model was proposed. By incrementally adding noise to the training samples, prior knowledge can be learned from the complex probability distribution. The acquired prior is then utilized as constraint in model-based iteration. The information of missing views can be gradually compensated by cyclic iteration to achieve high-quality reconstruction. The performance of the proposed method was evaluated with the circular phantom and *in vivo* experimental data. Experimental results demonstrate the outstanding effectiveness of the proposed method in limited-view cases. Notably, the proposed method exhibits excellent performance in limited-view case of 70° compared with traditional method. It achieves a remarkable improvement of 203% in PSNR and 48% in SSIM for the circular phantom experimental data, and an enhancement of 81% in PSNR and 65% in SSIM for *in vivo* experimental data, respectively. The proposed method has capability of reconstructing PAT images in extremely limited-view cases, which will further expand the application in clinical scenarios.

1. Introduction

Photoacoustic tomography (PAT) is an innovative hybrid medical imaging technology characterized by high penetration and rich contrast, and can achieve precise imaging of biological tissue structures at different spatial scales [1,2]. PAT has excellent application prospects in the fields of brain imaging [3], tumor detection [4], trauma assessment [5] and vascular structure imaging [6]. In PAT, biological tissue is irradiated with pulsed laser light. After biological tissue absorbs light energy, the thermoelastic expansion occurs, resulting in the emission of photoacoustic waves. Then, photoacoustic waves are detected by the ultrasonic transducers which is for the reconstruction of the biological tissue image [7,8]. The traditional reconstruction methods in PAT comprise delay-and-sum (DAS) [9], time-reversal [10] and filtered back-projection algorithms [11]. In PAT, the accurate reconstruction of biological tissue images theoretically requires meeting the conditions for full-view signal acquisition [12,13]. However, due to limitations in detector layout and sampling density, ultrasonic transducers can only collect photoacoustic signals within limited-view cases in practice [14,15]. The traditional methods are unable to recover the anticipated boundary information of the target biological tissues, the images reconstructed by these methods display severe artifacts and a noticeable decline in image quality. Therefore, how to use limited-view detection data to achieve high-quality reconstruction of biological tissue images has always been an urgent problem for PAT.

A series of studies have been conducted to solve this problem. Several iterative reconstruction algorithms have been proposed to improve image quality and reduce artifacts within limited-view cases [16–20]. These methods commonly require obtaining prior information regarding the acoustically inhomogeneous properties of biological tissues [19]. Yet, precise prior information in experiments is difficult to acquire due to the complexity and accuracy of calculations. In the field of biomedical imaging, post-processing methods based on deep learning stand out as the most prevalent approach recently [21–25]. In the image post-processing reconstruction, the convolutional neural network (CNN) is typically applied to eliminate artifacts and improve image quality.

* Corresponding author.
E-mail address: songxianlin@ncu.edu.cn (X. Song).
[1] contributed equally to this work

https://doi.org/10.1016/j.pacs.2024.100623
Received 28 January 2024; Received in revised form 11 May 2024; Accepted 17 May 2024
Available online 18 May 2024

Guan *et al.* designed a novel network called Pixel-DL to improve the quality and reduce artifacts [22]. Tong *et al.* proposed a deep learning approach named FPnet that integrates training strategy and data pre-processing to further enhance the quality of image [23]. Nevertheless, post-processing reconstruction demands extensive paired datasets for training. The acquisition of the intricate paired datasets in practice is challenging.

In recent years, new generative models such as variational autoencoder (VAE) [26], generative adversarial network (GAN) [27] and denoising diffusion probabilistic model (DDPM) [28] have attracted great attention in the biomedical field because of their powerful generative capabilities. Lu *et al.* put forwarded a novel hybrid data-driven deep learning method, denoted as LV-GAN, leveraging GAN to improve the quality of the limited-view PAT images [29]. GANs generate high-quality images by optimizing both the generator and discriminator. However, the training process of GANs is typically more unstable, requiring a large amount of data, and demanding high quality and diversity of data [30]. The diffusion model learns the data distribution and acquire valuable prior information directly through maximum likelihood estimation, enabling more effective utilization of limited data for training [31,32]. Chung *et al.* designed a network framework based on diffusion model to address inverse problems in the field of MRI imaging [33]. In preliminary work [34], a sparse-view reconstruction approach for PAT based on the diffusion model was proposed to reduce the image artifacts, achieving high-quality reconstruction under the 32 projections.

Inspired by this, this work proposes a limited-view photoacoustic image reconstruction scheme that combines model-based iterative reconstruction algorithm and score-based diffusion model. The prior information for image reconstruction is acquired through diffusion model, serving as a constraint on the data fidelity term within iterative reconstruction methods, contributing to improve reconstruction outcomes. In previous work [35], the reconstruction experiments were conducted on blood vessel simulation data, achieving successful reconstruction in limited-view case of 90°. Nevertheless, the previous work lacked validation with experimental data and comparisons with other mainstream methods. To address this issue, the effectiveness of this method was further verified in this paper. The performance of the proposed method was verified on circular phantom and *in vivo* experimental data, with comparisons made against other methods. Experimental results show that the method has outstanding performance within extremely limited-view cases. The impacts of this work are summarized as follows:

- A limited-view PAT reconstruction method integrating diffusion model and model-based iteration was proposed, achieving high-quality reconstruction in extremely limited-view cases.
- The employed score-based diffusion model is an unsupervised network, which can be trained on a minimal and unpaired dataset, without the need for repetitive training in different limited-view cases.
- Breast cancer detection [36] and brain imaging [37] often require photoacoustic signals acquisition in limited-view cases [38], and the proposed method can ameliorate the imaging quality degradation caused by limited-view data.

2. Principles and methods

2.1. The model-based iterative method of PAT

In PAT, the initial sound pressure distribution of the biological tissue is reconstructed by the detected photoacoustic signal, so as to obtain the information on the structure and function of the biological tissue. Employing the Green's function method, the relationship between the sound pressure distribution of target biological tissue and the propagation of photoacoustic signals can be described as Eq. (1):

$$p\left(r,t\right)=\frac{1}{4\pi c^2}\frac{\partial}{\partial t}\left[\frac{1}{ct}\int d\acute{r}p_0\left(\acute{r}\right)\delta\left(t-\frac{|r-\acute{r}|}{c}\right)\right] \tag{1}$$

where$p(r,t)$is the sound pressure distribution at position r and time t, c signifies the speed of sound and$p_0$represents the initial pressure distribution of the biological tissue. Eq. (1) can be equivalently transformed into the linear procedure shown in Eq. (2) [39]:

$$y=Ax \tag{2}$$

where y is the limited-view photoacoustic signal, which is detected by the ultrasonic transducer. x signifies the initial sound pressure p_0 of the biological tissue. And the linear operator A is the propagation process of the photoacoustic signal. The limited-view PAT is to use the limited-view photoacoustic signal y to calculate the initial sound pressure x. This problem is also called the inverse problem of acoustics, which can be solved by the least-square method, as shown in Eq. (3):

$$x=\underset{x}{\operatorname{argmin}}\|Ax-y\|_2^2 \tag{3}$$

where $\|Ax-y\|_2^2$is the data fidelity term. Owing to the existence of noise and the incomplete collection of data, this inverse acoustic problem is ill-posed, and the reconstructed photoacoustic image has serious artifacts and distortions [40]. Therefore, it is necessary to add a regularization term or constraint term to reduce the ill-posedness and improve the reconstruction quality of the image. These constraints originate from the inherent features of the image, termed as prior information. The prior information can be introduced into Eq. (3) as a regularization term, thereby converting the optimization problem into Eq. (4).

$$x=\underset{x}{\operatorname{argmin}}\|Ax-y\|_2^2+\lambda R(x) \tag{4}$$

where $R(x)$ represents the regularization term, andλ represents the regularization parameter. The regularization term can solve the problem of non-uniqueness and instability of the inverse reconstruction process to a certain extent [41]. Universal regularization methods include total variational regularization [42] and Tikhonov regularization [43]. Conventional regularization terms solely grasp basic structural information within the image, rendering the acquisition of specific feature details challenging. Diffusion model can learn abundant prior knowledge as regularization terms to constrain the solution of optimization problems, thereby enhancing the quality of the images.

2.2. Diffusion model

The diffusion model fits given samples by utilizing probability distributions. Probability distribution is generally represented applying a score function (the gradient of the logarithmic probability density). Diffusion model can learn the score function $\nabla_x \log p_t(x)$ to generate new high-quality samples from noise. Fig. 1 illustrates two processes of the diffusion model: the forward diffusion (stochastic differential equation, SDE) and reverse diffusion. During the forward diffusion, the model slowly adds Gaussian noise into the training samples. This process can smoothly transform the complex probability distribution of the training samples into known prior information. The diffusion process can be represented as $\{x_t\}_{t=0}^{T}$, where $t\in[0,T]$ is a variable of time. $x_0\sim p_{data}$, p_{data} represents the initial probability distribution of the training sample without noise perturbation. $x_T\sim p_T$, $p_T$signifies the probability distribution of the training sample after undergoing noise perturbation for a duration of T time. The forward diffusion process can be modeled as shown in Eq. (5):

$$dx=f(x,t)dt+g(t)dw \tag{5}$$

where w is a standard Brownian motion,$f(x,t)\in\mathbb{R}$ represents the drift coefficient and $g(t)\in\mathbb{R}$ is the diffusion coefficient. The reverse SDE is

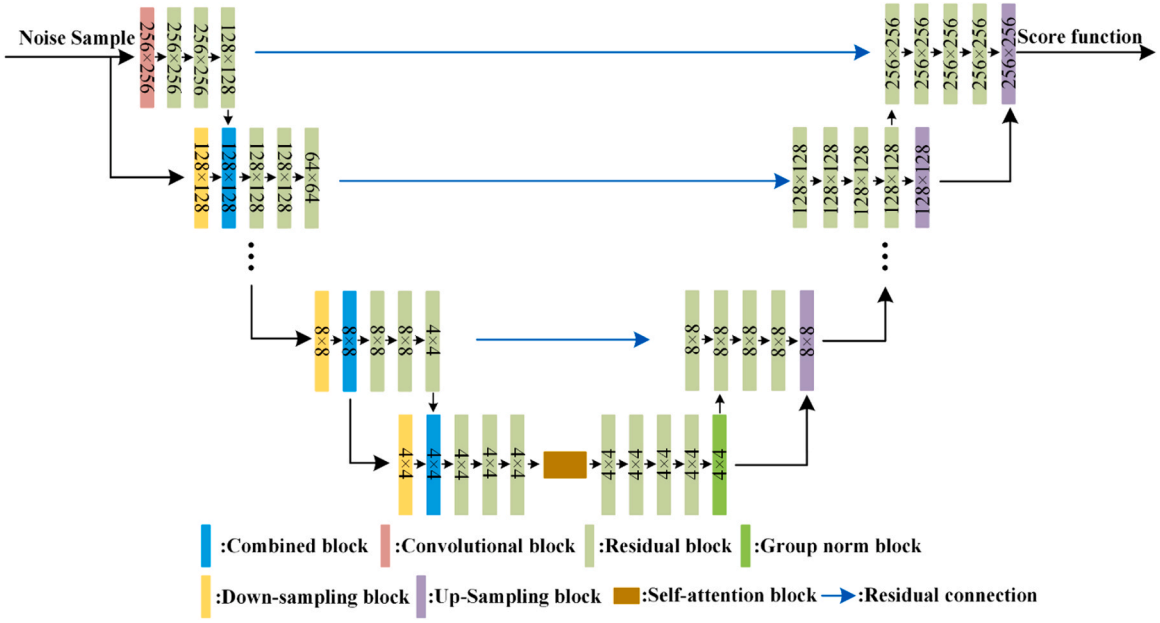


Fig. 1. The score network architecture.

defined as a process that acquires samples from the noise, and its solution can be determined using the score function. The reverse SDE is illustrated in Eq. (6):

$$dx = \left[f\left(x,t\right) - g(t)^2\nabla_x \log p_t(x)\right]dt + g(t)d\overline{w} \tag{6}$$

where$\nabla_x \log p_t(x)$ signifies the score function of the data and $\overline{w}$is the reverse Brownian motion. Utilizing different $f(x,t)$ and $g(t)$ can construct distinct categories of SDE. Empirically, using variance exploding (VE) SDE can yields higher-quality samples. Hence, this study focuses on diffusion model based on VE-SDE, with the$f(x,t)$ and $g(t)$ functions as shown in Eq. (7):

$$f\left(x,t\right) = 0, g(t) = \sqrt{\frac{d[\sigma^2(t)]}{dt}} \tag{7}$$

where $\sigma(t) > 0$ represents a monotonically increasing Gaussian noise function. The reverse diffusion process of VE-SDE can be described as Eq. (8):

$$dx = -d\left[\sigma^2(t)\right]\nabla_x \log p_t(x) + \sqrt{\frac{d[\sigma^2(t)]}{dt}}d\overline{w} \tag{8}$$

However, the score function $\nabla_x \log p_t(x)$ of given data cannot be directly obtained. The estimation of score function can be achieved by training a time-dependent score network $S_\theta(x_t,\ t)$. The unknown $\nabla_x \log p_t(x)$can be substituted with $\nabla_{x_t} \log p_t(x_t|x_0)$utilizing the denoising score matching. The $\nabla_{x_t} \log p_t(x_t|x_0)$represents the Gaussian perturbation kernel centered on x_0. The score network architecture mainly consists of convolutional blocks, residual blocks and self-attention blocks, as shown in Fig. 1. The basic architecture of the score network is based on the U-Net [44], featuring an encoder-decoder framework. The encoder part is responsible for feature extraction, and the decoder part is involved in image resolution restoration. Self-attention blocks are incorporated at the junction between the encoder and decoder to enhance the ability of network to perceive crucial features [45,46]. The purpose is to enable the network to capture rich feature information during the encoder stage and effectively convey this information to generate results during the decoder stage. This aims to enhance the performance of network in various image processing tasks. The residual blocks of the network are identical to those of BigGAN [47].

The training procedure of the score network is depicted in the upper half part of Fig. 2. The objective function for training the score network can be represented as Eq. (9):

$$\theta^* = \arg\min_\theta E_t\left\{\lambda(t)E_{x_0}E_{x_t|x_0}\left[\|S_\theta(x_t,t) - \nabla_{x_t}\log p_t(x_t|x_0)\|_2^2\right]\right\} \tag{9}$$

where E is the expectation,$E_t\left\{\lambda(t)E_{x_0}E_{x_t|x_0}\left[\|S_\theta(x_t,t) - \nabla_{x_t}\log p_t(x_t|x_0)\|_2^2\right]\right\}$can be considered as the loss function, and $\lambda(t)$ is a time-related weight function. The trained score network $S_\theta(x_t,t) \simeq \nabla_x \log p_t(x_t)$ can be substituted into the reverse SDE for a solution. As shown in Eq. (10):

$$dx = -d\left[\sigma^2(t)\right]S_\theta\left(x_t,t\right) + \sqrt{\frac{d[\sigma^2(t)]}{dt}}d\overline{w} \tag{10}$$

2.3. The PAT reconstruction based on diffusion model

The classical approach to PAT reconstruction is the solution of Eq. (4). The data fidelity term in Eq. (4) can be solved using the gradient descent (GD) method, as shown in Eq. (11):

$$x_{i-1} = x_i - \alpha A^*(Ax_i - y) \tag{11}$$

where α signifies the iteration step size, A^* represents the adjoint operator of A, and $A^*(Ax_i - y)$ is the gradient of the data fidelity term. Provided that the regularization term is differentiable, the iterative solution of Eq. (4) can be further expressed as Eq. (12):

$$x_{i-1} = x_i - \alpha\left[A^*\left(Ax_i - y\right) + \lambda\frac{\partial R(x_i)}{\partial x}\right] \tag{12}$$

In the reconstruction process, as shown in the lower half part of Fig. 2, the data fidelity term and the regularization term in Eq. (4) are decoupled. Following the rotational solving principle [48,49], the optimal solution is achieved through alternating updates of two sub-problems. This process can be expressed as Eq. (13).

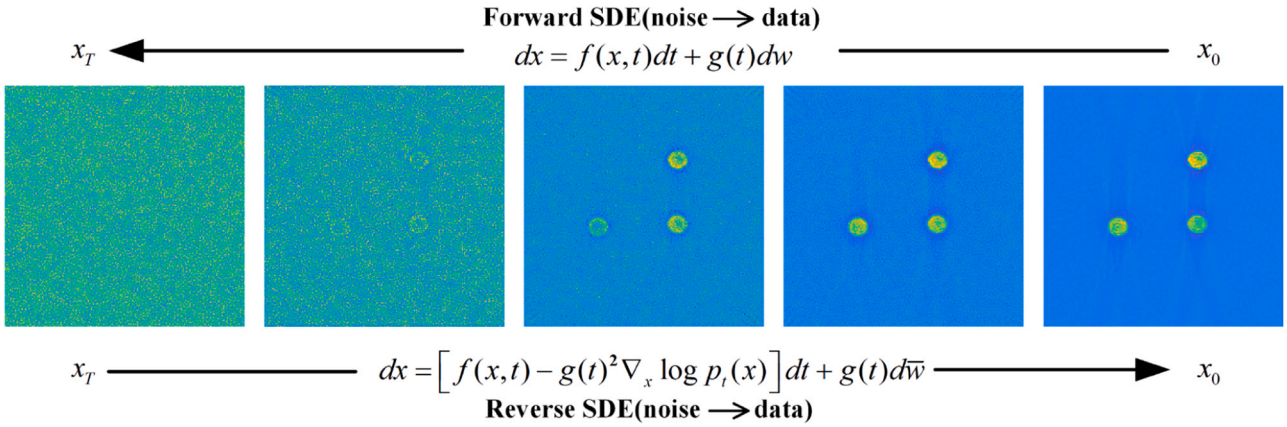


Fig. 2. The forward and reverse diffusion process of the diffusion model.

$$\begin{cases} \hat{x}_i = x_i + \beta \dfrac{\partial R(x_i)}{\partial x} = \mathrm{G}(x_i) \\ x_{i-1} = \hat{x}_i - \alpha A^*(A\hat{x}_i - y) \end{cases} \tag{13}$$

where $\mathrm{G}(x_i)$ represents the output of network containing prior information in the first equation. x_i represents the input of diffusion model, $\beta \frac{\partial R(x_i)}{\partial x}$ signifies the gradient of the regularization term, β is a parameter associated with the regularization term. $\hat{x}_i$ is an intermediary variable, essentially representing the output of the diffusion model, and serving as the input to the gradient descent. The second equation symbolizes the iteration solution to the data fidelity using the GD method. For the first equation, it can be resolved by the reverse SDE of Eq. (10). The Euler discretization method can be used to predict the numerically solve of the reverse SDE. With the aim to correct the error in the discretization evolution of the reverse SDE, the Langevin dynamics Markov chain Monte Carlo method [50] is employed as a corrector. This process is termed as prediction-correction (PC) sampling. After settling the reverse SDE and performing correction algorithm, a fidelity operation is applied as shown in the second equation of Eq. (13). Thus, the algorithm of the limited-view PAT reconstruction based on diffusion model can be divided into two components: the PC sampling and the fidelity operation. In the PC sampling component, the objective is to generate the image containing prior information from the learned probability distribution. It can be addressed by Eq. (14) and Eq. (15):

$$\tilde{x}_i = x_i + \left(\sigma_{i+1}^2 - \sigma_i^2\right) S_\theta\left(x_i, \sigma_{i+1}\right) + \sqrt{\sigma_{i+1}^2 - \sigma_i^2}\, z \tag{14}$$

where σ_i represents the noise scale, i denotes the count of iteration, and z is zero-mean Gaussian white noise. After the prediction phase, the correction algorithm is employed to correct the error in the discretization evolution, as shown in Eq. (15):

$$\hat{x}_i = \tilde{x}_i + \varepsilon_i S_\theta\left(\tilde{x}_i, \sigma_{i+1}\right) + \sqrt{2\varepsilon_i} z \tag{15}$$

After the PC sampling, fidelity operation is enforced on the results as shown in Eq. (16), which can effectively constrain the generated samples.

$$x_{i-1} = \hat{x}_i - \alpha A^*(A\hat{x}_i - y) \tag{16}$$

The pseudocode of the proposed reconstruction algorithm is presented in Process 1. This process comprises two processes: the training process and the reconstruction process. In the training process, the network learns the data distribution of the given samples. The

Process 1

Training for prior learning

Dataset: Full-view dataset x

Estimation of score function: $S_\theta(x_t, t)$; $\nabla_x \log p_t(x_t)$

Reconstruction

Setting: ε_i : Noise iteration step, N: Number of discretization steps for the reverse SDE, M: Number of correctors steps

1: $x_N \quad (0, \sigma_{\max}^2 I)$

2: For i =N-1 to **0** **(Outer loop)**

3: $\tilde{x}_i = x_i + (\sigma_{i+1}^2 - \sigma_i^2) S_\theta(x_i, \sigma_{i+1}) + \sqrt{\sigma_{i+1}^2 - \sigma_i^2}\, z$ **Prediction**

4: **For** j =1 to M **(Inner loop)**

5: $\hat{x}_i \leftarrow \tilde{x}_{i,j} + \varepsilon_i S_\theta(\tilde{x}_{i,j}, \sigma_{i+1}) + \sqrt{2\varepsilon_i} z$ **Correction**

6: **End**

7: $x_{i-1} = \hat{x}_i - \alpha A^*(A\hat{x}_i - y)$ **Gradient descent**

8: End

Output x_0

reconstruction process consists of two nested loops. In the outer loop, the model attains a forecast of samples using the learned data distribution from training. In the inner loop, the prediction is corrected using the correction method. After each correction, fidelity operation is enforced on the reconstructed results to achieve an optimal solution.

2.4. Experimental parameter setting and dataset acquisition

To accurately simulate the forward propagation process of photoacoustic signals in heterogeneous media, the virtual PAT based on k-Wave [51] was constructed. Fig. 3 shows the virtual PAT. The virtual PAT region has a size of 50 mm×50 mm. This region is divided into the grid of 440×440 pixels to discretize the medium. The ultrasonic transducer array is placed in an arc with a radius of 21.6 mm around the imaging object. 70, 90, 120 and 180 ultrasonic transducer array elements cover 70°, 90°,120° and 180° around the target issue, respectively. The bandwidth of the ultrasonic transducers is 66%, and the center frequency is 2.25 MHz. The speed of sound for the photoacoustic signal propagation is taken as 1500 m/s, and the medium surrounding the imaging object is water with a density of 1000 kg/m^3.

The datasets used for training the network consists of a circular phantom dataset and an *in vivo* experimental dataset of mice's abdomens. The circular phantom dataset is derived from 469 full-view scan results (512 detectors) of a circular phantom dataset [25]. 1600 images with a size of 256×256 are collected through data augmentation with rotation and cropping. The training and test sets are allocated at a ratio of 8:1, comprising 1422 images for training and 178 images for testing. The *in vivo* experimental dataset is obtained from [25], and it is sourced from 274 full-view (512 detectors) scan images. These images are subjected to rotation enhancement, culminating in a collection of 800 images with a size of 256×256. There are 711 images for training and 89 images for testing. During the training phase, the diffusion model learns the data distribution of the full-view (512 detectors) PAT images by gradually adding Gaussian noise. During the reconstruction phase, the detected time series is employed in the model-based iterative solution.

The Adaptive Moment Estimation (Adam) optimizer is used in the training of the diffusion model to minimize the loss function, with a learning rate of 2×10^{-4}. The noise scale in the SDE process ranges from 0.01 to 300. The training batch size is set to 1. One checkpoint is saved every 10,000 epochs completed, and twenty checkpoints are obtained in the training. The number of iterations of the model is taken as 900, and the size of the input and output images is set to 256×256. The network is implemented in Python version 3.8.15 and the PyTorch framework version 1.7.1. The experiments are executed on a workstation computer equipped with GeForce RTX 3060Ti.

2.5. Baseline methods

The DAS, GAN, Tikhonov, and U-Net methods are selected for comparison with the proposed method in this study. The GAN method is derived from Cycle-GAN [52]. Cycle-GAN can achieve mutual transformation between two different types of images with unpaired training set, and its outstanding performance has been validated in PAT images [53]. The Tikhonov method is the GD method with Tikhonov regularization. The U-Net network employed in this work was derived from [25]. There are two main components in the U-Net network: encoder and decoder. The encoder module is responsible for extracting features from input data, while the decoder generates corresponding outputs based on the features extracted by the encoder. Both parts consist of consecutive convolutional layers, with the channel range of convolutional layers varying from 64 to 1024.

For the training and test sets of U-Net, a distribution scheme consistent with the diffusion model was adopted. During the training phase, limited-view PAT images are fed into the network, while full-view PAT images serve as the gold standard. The Adam optimization algorithm is utilized to minimize the loss, with an initial learning rate set to 0.005. The epoch was set to 200. The U-Net training loss function was characterized by the average L1 norm between the predicted image and the reference full-view image.

3. Results

3.1. Results on circular phantom experimental data

The feasibility of the proposed method is evaluated on circular phantom data. Figs. 5(a), 5(b) and 5(c) illustrate the iterative reconstruction process of circular phantom using the proposed method, GD method and Tikhonov method (GD method with Tikhonov regularization) in limited-view case of 70°, respectively. The white number at the lower-right corner of the figure is peak signal to noise ratio (PSNR), and

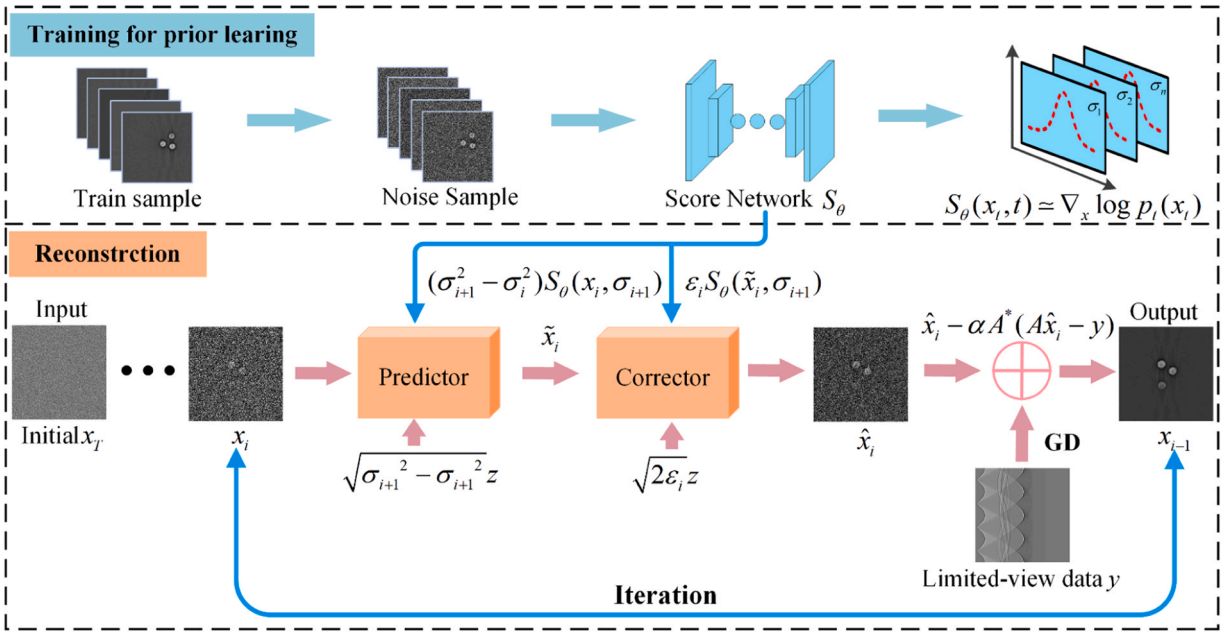


Fig. 3. Reconstruction flow chart of the limited-view PAT based on diffusion model. GD: gradient descent.

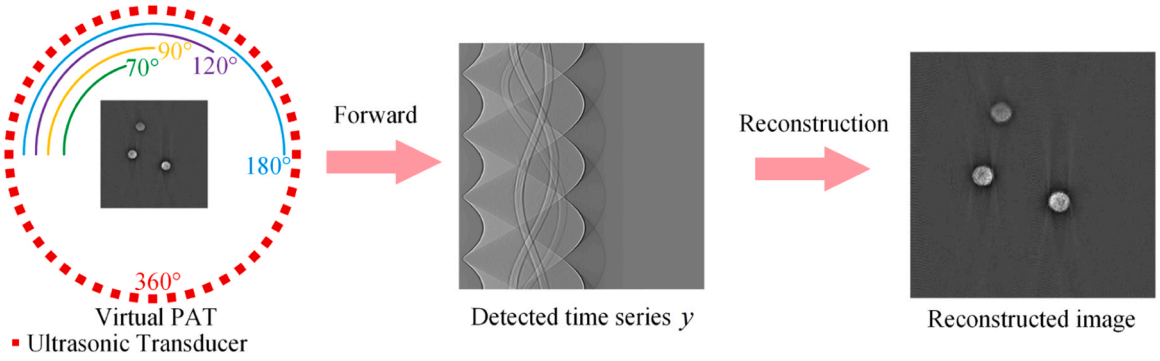


Fig. 4. Reconstruction of the PAT images using virtual PAT.

the yellow number is similarity index (SSIM). The iterative reconstruction starts with the same Gaussian noise. Following 10 iterations, all methods restore relatively rough contour of the circular phantom. However, the contour of the circular phantom still exhibits conspicuous elongation and severe noise is still present at this stage (indicated by the white arrows in Figs. 5(a), 5(b) and 5(c)). After 100 iterations, the circular phantom reconstructed by the proposed method becomes visibly clearer. At the 900th iteration, the proposed method has essentially completed the reconstruction of the circular phantom, leading to a further improvement in image quality (indicated by yellow arrows in Fig. 5(a)). The reconstruction results of the GD method and Tikhonov method manifest significant noise, and the circular phantom is elongated and distorted (indicated by yellow arrows in Figs. 5(b) and 5(c)). Fig. 5(a1), 5(b1) and 5(c1) are the Fourier spectra corresponding to Figs. 5(a), 5(b) and 5(c), respectively. As the number of iterations increases, the information in Fourier space recovered by the proposed method also grows, suggesting an augmentation in details and more distinct delineations. The lower-left corner of Figs. 5(a), 5(b) and 5(c) show the corresponding error maps for the reconstruction results. Figs. 5(d) and 5(e) display the variations of PSNR and SSIM during the iterative process of all methods. For the proposed method, the PSNR and SSIM of the results rapidly increase before the 700th iteration. At the 700th iteration, the PSNR and SSIM of the results reach 28.26 dB and 0.95, respectively. With further iterations, the variation of PSNR and SSIM stabilize (additional information on the iteration process is available in Visualization 1). The experimental results demonstrate that the proposed method can effectively accelerate the iterative process and enhance image quality.

Fig. 6 presents the reconstruction results of circular phantom applying different methods in limited-view cases of 180°, 120°, 90° and 70°, respectively. Fig. 6(a1)-6(a4), 6(b1)-6(b4), 6(c1)-6(c4), 6(d1)-6(d4), 6(e1)-6(e4) and 6(f1)-6(f4) showcase the results using the DAS, GD, Tikhonov, U-Net, GAN and proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. The white number at the lower-left corner of the figure is PSNR, and the yellow number is SSIM. Fig. 6(a5)-6(f5) are the same ground truth, which is the PAT image reconstructed by DAS in full-view case (512 detectors). It can be noticed that the DAS method fails to reconstruct the complete contour of the circular phantom, resulting in significant distortion of the results. For the GD method, reconstructing high-quality images is quite challenging. As the detection angle decreases, the image quality deteriorates. The reconstruction results of the Tikhonov method show slight improvement compared to the GD method. However, in extremely limited-view cases, the reconstructed results still present significant distortions. In general, the results reconstructed by the U-Net and GAN method exhibit more complete contour of the circular phantom compared to the DAS, GD and Tikhonov method. However, distortions and artifacts persist in the result using the U-Net and GAN in limited-view cases of 90° and 70°. The results using the proposed method display a superior quality and clearer delineation in different limited-view cases. Even in exceedingly constrained limited-view cases (e.g., 70°-view), the results utilizing the proposed method have adequate details of the circular phantom.

Figs. 7(a) and 7(b) represent close-up images indicated by the red rectangles 1 and 2, respectively. Similarly, close-up images highlight the superiority of the proposed method. The reconstruction results of the proposed method display reduced artifacts and increased fidelity in details (as pointed out by the white arrows in Figs. 7(a) and 7(b)) compared to other methods. Figs. 7(c) and 7(d) depict intensity profiles along the dashed lines, revealing that the intensity profiles of the images reconstructed by the proposed method exhibit closer proximity to the ground truth. Fig. 8 shows the error maps between the reconstruction results and the ground truth, which indicate that the results using the proposed method exhibit smaller errors to the ground truth. In terms of quantitative image comparison, the proposed method attains a PSNR of 31.57 dB and an SSIM of 0.95 in limited-view cases of 70°, demonstrating enhancements of 21.16 dB (~203%) and 0.31 (~48%) over the DAS method, respectively. The proposed method achieved improvements of 3.33 dB and 0.03 over the U-Net method and 10.17 dB and 0.10 over the GAN method, respectively. This further validates the excellence and effectiveness of the proposed method in limited-view PAT reconstruction.

Fig. 9 illustrates the Fourier spectra transformed from the reconstructed results through the Fourier transform. Fig. 9(a1)-9(a4), 9(b1)-9(b4), 9(c1)-9(c4), 9(d1)-9(d4), 9(e1)-9(e4) and 9(f1)-9(f4) depict the Fourier spectra corresponding to the reconstruction results of the DAS, GD, Tikhonov, U-Net, GAN and proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. It can be spotted that the Fourier spectra of the ground truth (Fig. 9(a5)-9(f5)) is uniformly distributed in the Fourier space in full-view case. The Fourier spectra corresponding to the results of the DAS, GD and Tikhonov method only contain information at angles covered by the ultrasonic transducer, with missing information at other angles (as pointed out by the white arrows in Fig. 9(a2), 9(a3) and 9(a4)) which leads to distorted contours and blurred backgrounds. Although the U-Net method can supplement some information, there is still great information loss compared to the ground truth. The GAN method demonstrates the ability to augment more high-frequency information in contrast to the U-Net method. The most information in the Fourier space can be recovered by the proposed method in different limited-view cases compared with other methods, suggesting that the quality of the reconstruction results can be greatly improved.

To better showcase the superiority of the proposed method, five images from the circular phantom test set are selected for the reconstruction experiments. Table 1 shows the mean and standard deviation for PSNR, SSIM, feature similarity index measure (FSIM) [54] and mean squared error (MSE) of the reconstruction results. The FSIM employs feature similarity for quality assessment, with a predominant emphasis on the structural and textural characteristics of images. In limited-view case of 70°, the proposed method attains an average PSNR of 31.33 dB and an average SSIM of 0.94, with an improvement of 20.43 dB (~186%) and 0.31 (~ 49%) compared with the DAS method, respectively.

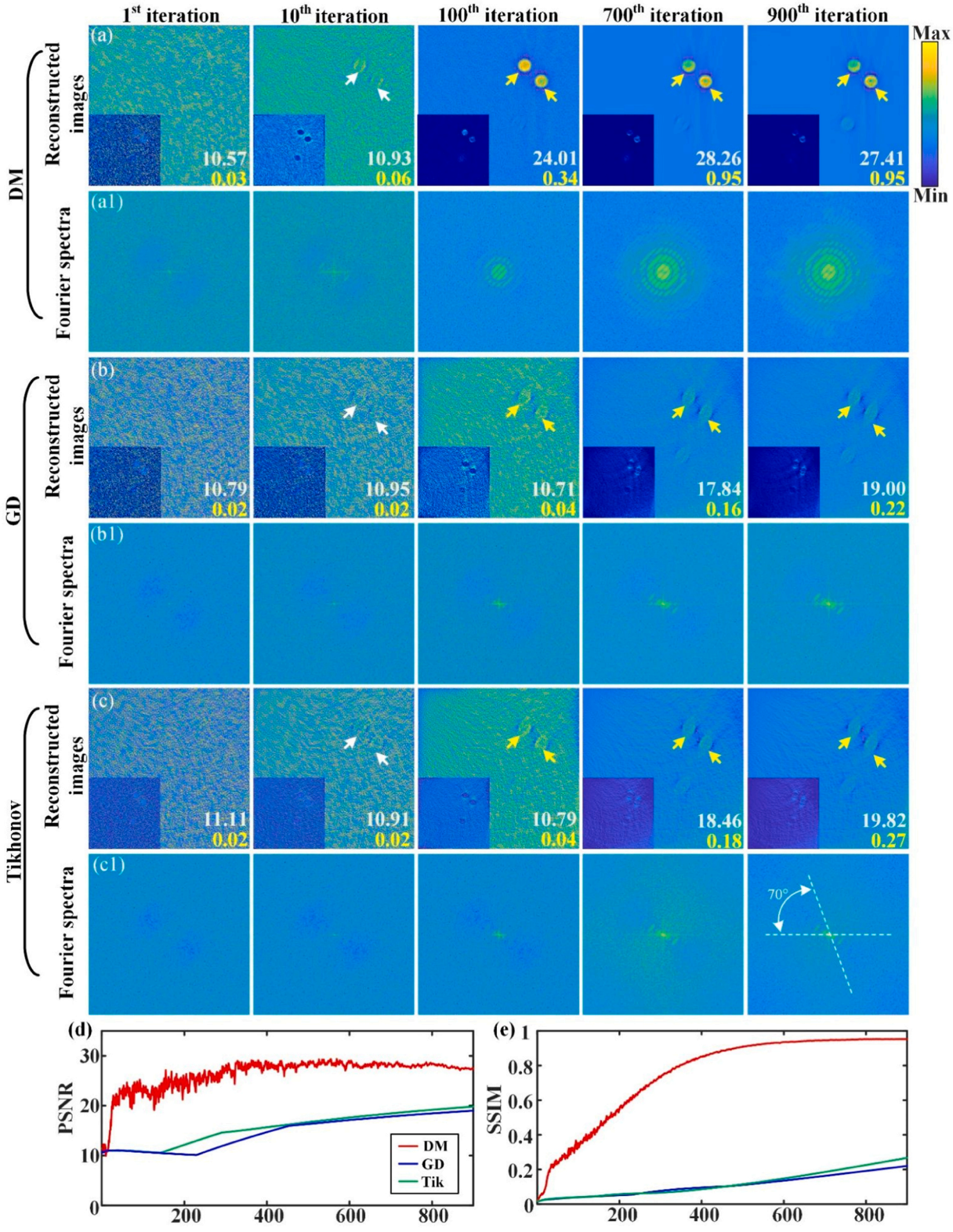


Fig. 5. The reconstruction process of the circular phantom in limited-view case of 70°. (a) displays the reconstruction results using the proposed method, with white numbers indicating PSNR and yellow numbers indicating SSIM. (a1) is the Fourier spectra corresponding to (a). (b) shows the reconstruction results using the GD method. (b1) is the Fourier spectra corresponding to (b). (c) displays the reconstruction results using the Tikhonov method. (c1) is the Fourier spectra corresponding to (c). The lower-left corner of (a), (b) and (c) show the corresponding error maps for the reconstruction results. (d) and (e) depict the variations in PSNR and SSIM during the iterations. The white arrows in (a), (b) and (c) denote the circular phantom which is not completely reconstructed. The yellow arrows in (a), (b) and (c) indicate the circular phantom contours reconstructed by the DM, GD and Tikhonov method at the 100th, 700th and 900th iterations, respectively. DM, diffusion model. GD, gradient descent. Tik, the Tikhonov method.

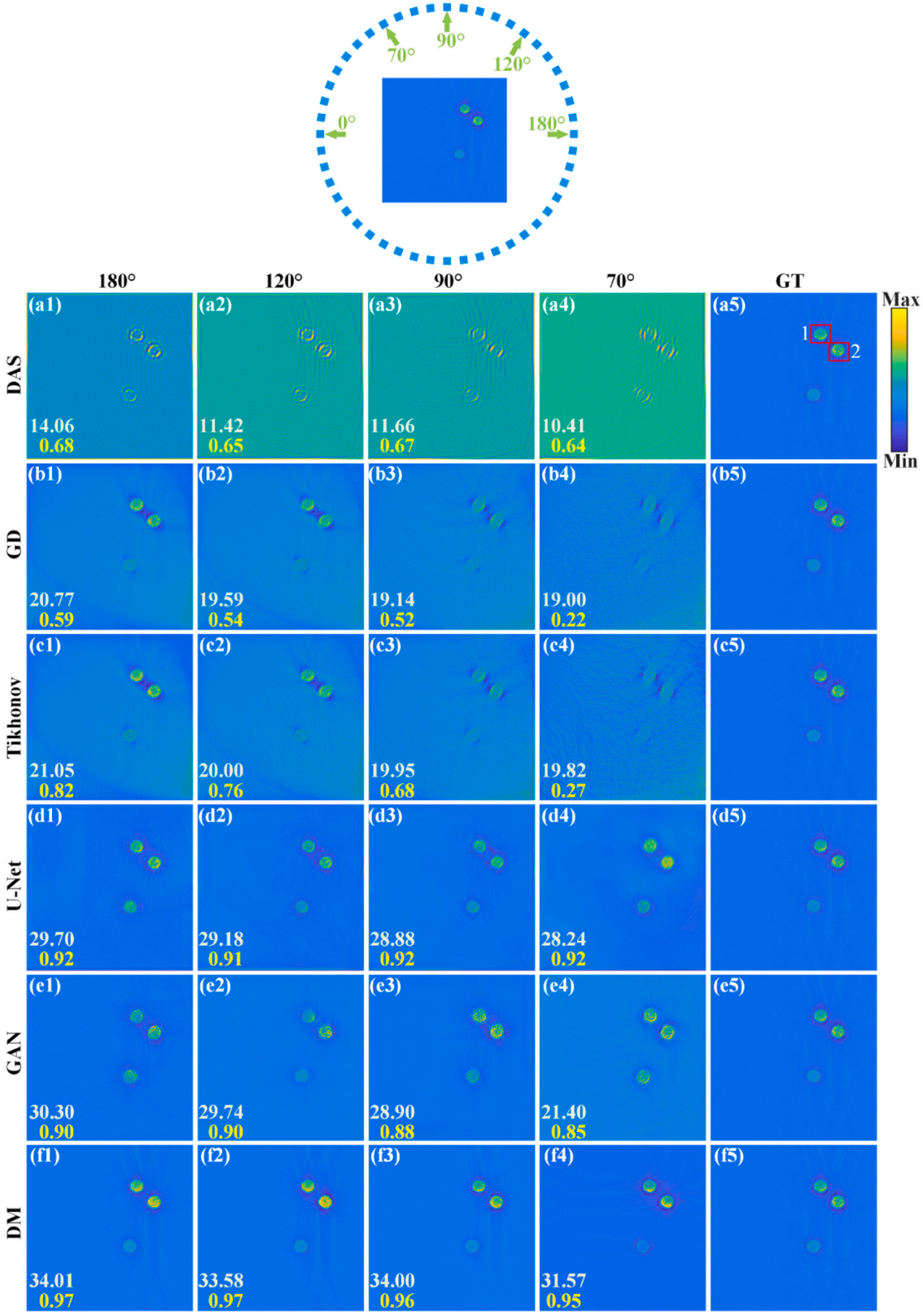


Fig. 6. The reconstruction results of circular phantom. (a1)-(a4) represent the results of the DAS method in limited-view cases of 180°, 120°, 90° and 70°, respectively. The white number at the lower-left corner of the figure is PSNR, and the yellow number is SSIM. (b1)-(b4) are the results applying the GD method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (c1)-(c4) show the results applying the Tikhonov method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (d1)-(d4) display the results using the U-Net method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (e1)-(e4) are the results applying the GAN method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (f1)-(f4) represent the results of the proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (a5)-(f5) are the same ground truth. DM, diffusion model. GD, gradient descent. GT, ground truth.

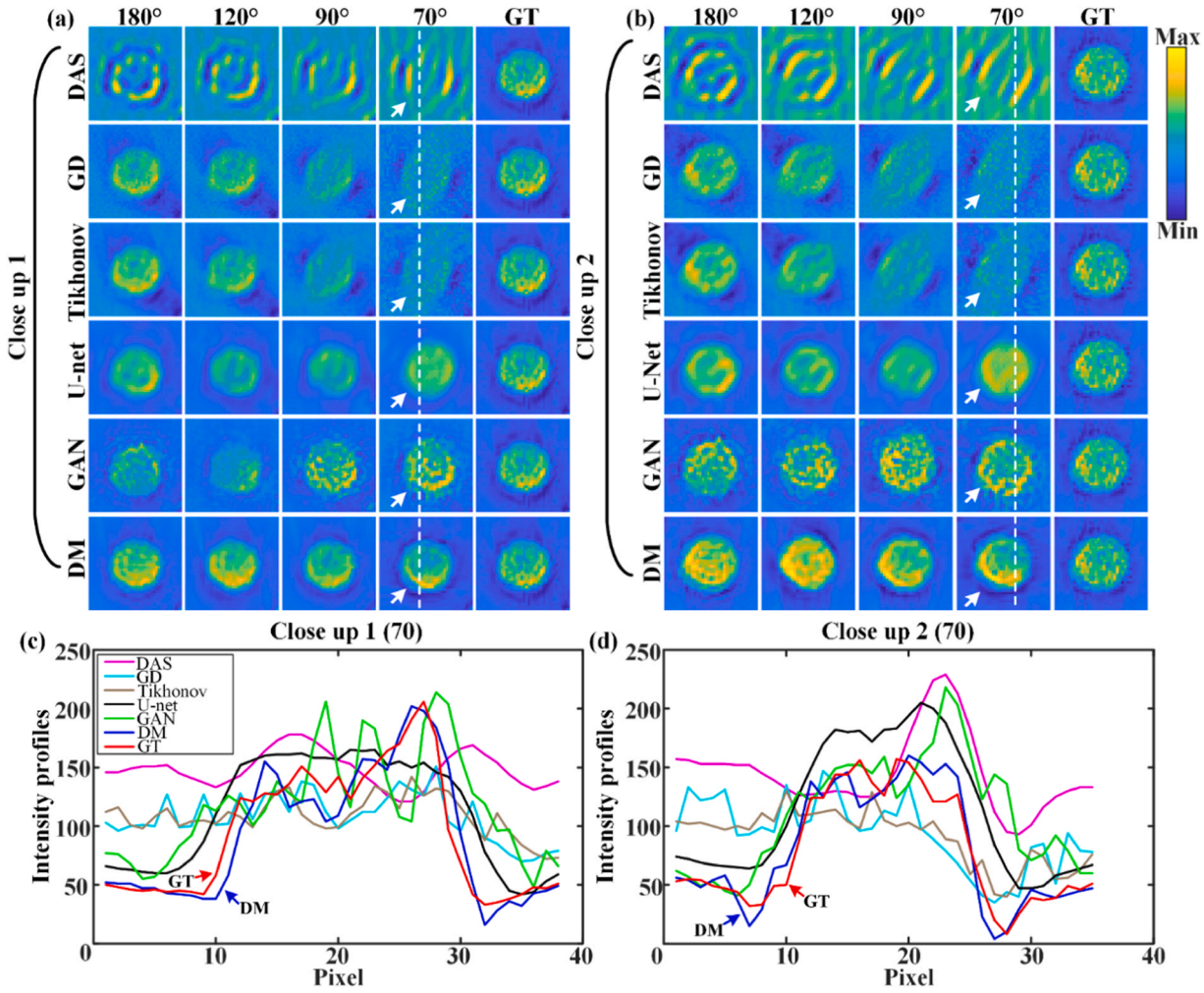


Fig. 7. The close-up images and intensity profiles along the dashed lines of the circular phantom. (a) and (b) represent the close-up images indicated by the red rectangles 1 and 2, respectively. (c)-(d) show the intensity profiles along the dashed lines in (a) and (b), respectively. The white arrows indicate a reduction in artifacts and an improvement in detail fidelity compared to other methods in the reconstruction results of the proposed method. DM, diffusion model. GD, gradient descent. GT, ground truth.

In limited-view case of 90°, the reconstructed results of the proposed method exhibit average improvement of 5.37 dB and 0.04 compared to the U-Net method, respectively. This further demonstrates the superior and valid performance of the proposed method in PAT reconstruction within limited-view cases.

3.2. Results on in vivo experimental data

To further confirm the application potential of the proposed method in the limited-view PAT, experiments were subsequently conducted on *in vivo* experimental data of mice's abdomens. Fig. 10 shows the reconstruction result using the DAS method, U-Net method and the proposed method, respectively. Figs. 10(a)-10(d) display the results of the DAS method in limited-view cases of 180°, 120°, 90° and 70°, respectively. It can be observed that the reconstruction results using the DAS method have severe artifacts, low image contrast and the physiological structure of the mice's abdomens is significantly distorted. Figs. 10(e), 10(j) and 10(o) are the same ground truth, which is the PAT image reconstructed by DAS within full-view case (512 detectors). Figs. 10(f)-10(i) demonstrate the results of the U-Net method in limited-view cases of 180°, 120°, 90° and 70°, respectively. The U-Net method improves the image contrast to some extent. However, there is conspicuous detail loss in the results obtained by the U-Net method. Figs. 10(k)-10(n) show the results of the proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. The results employing the proposed method reveal a substantial enhancement in quality compared with DAS and U-Net method in different limited-view cases. The visibility of the structure is remarkably improved employing the proposed method, especially in extremely confined limited-view cases (e.g., 70°-view, additional information on the iteration process is available in Visualization 1). Figs. 10(p) and 10(q) represent the close-up images indicated by the red rectangles 1 and 2, respectively. It is evident that the proposed method can precisely reconstruct the details of the mice's abdomens compared with the U-Net method (as pointed out by the white arrows). Intensity profiles in Figs. 10(r) and 10(s) also indicate that the intensity profiles of the results reconstructed by the proposed method closely resemble the ground truth.

Figs. 11(a)-11(o) illustrate the error maps between the results and the ground truth. The error maps demonstrate that the reconstruction images using the proposed method have smaller discrepancies and more proximate resemblance from the ground truth. Table 2 shows the PSNR, SSIM, FSIM and MSE of the results. Quantitative comparison further supports the superiority of the proposed method. In limited-view case of 180°, the proposed method provides higher PSNR and SSIM ability. There is an enhancement of 1.22 dB in PSNR and 0.15 of SSIM compared with the U-Net method, respectively. In limited-view case of 90°, the proposed method presents the PSNR of 29.18 dB and the SSIM of 0.80, with an improvement of 11.37 dB (~64%) and 0.26 (~48%) compared

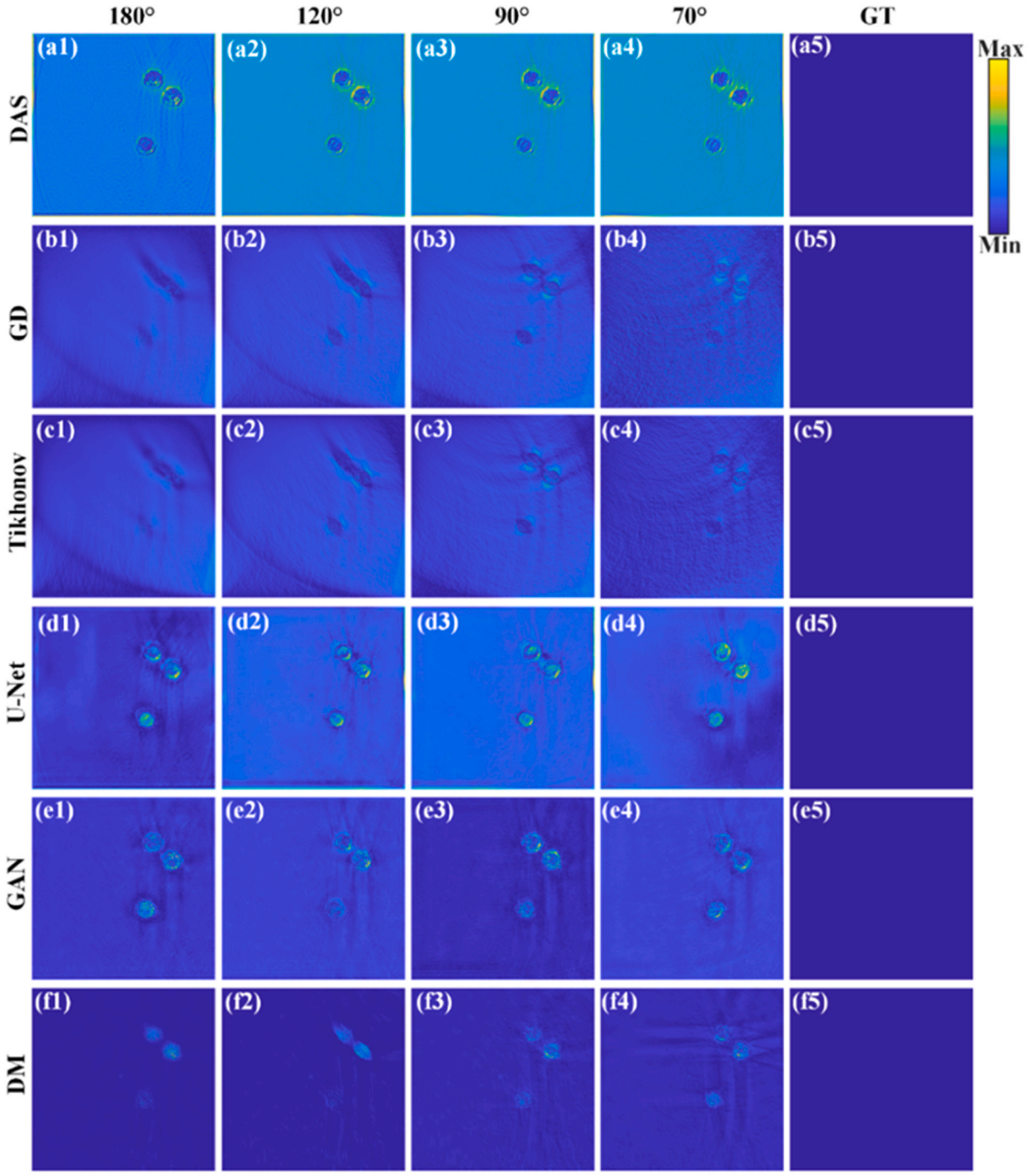


Fig. 8. Error maps obtained from the reconstruction results of the circular phantom data. (a1)-(a4) represent the error maps of the DAS method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (b1)-(b4) are the error maps applying the GD method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (c1)-(c4) show the error maps applying the Tikhonov method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (d1)-(d4) display the error maps using the U-Net method in limited-view casas of 180°, 120°, 90° and 70°, respectively. (e1)-(e4) are the error maps applying the GAN method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (f1)-(f4) represent the error maps of the proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (a5)-(f5) are the same ground truth. DM, diffusion model. GD, gradient descent. GT, ground truth.

with the DAS method, respectively. The experimental results indicate the proposed method possesses exceptional performance on boosting image contrast and eradicating artifacts in extremely restrict limited-view cases.

4. Conclusion and discussion

In summary, to alleviate the influence of artifacts and distortions in the reconstruction of the limited-view PAT, this manuscript introduces a high-quality reconstruction method based on the score-based diffusion model. Within the training phase, the model learns the data distribution of samples by gradually adding noise into existing samples. During the reconstruction phase, the learned prior knowledge of the diffusion model serves as the regularization term of the traditional iterative reconstruction method, enabling to facilitate the generation of high-quality PAT images. The performance of the proposed method was

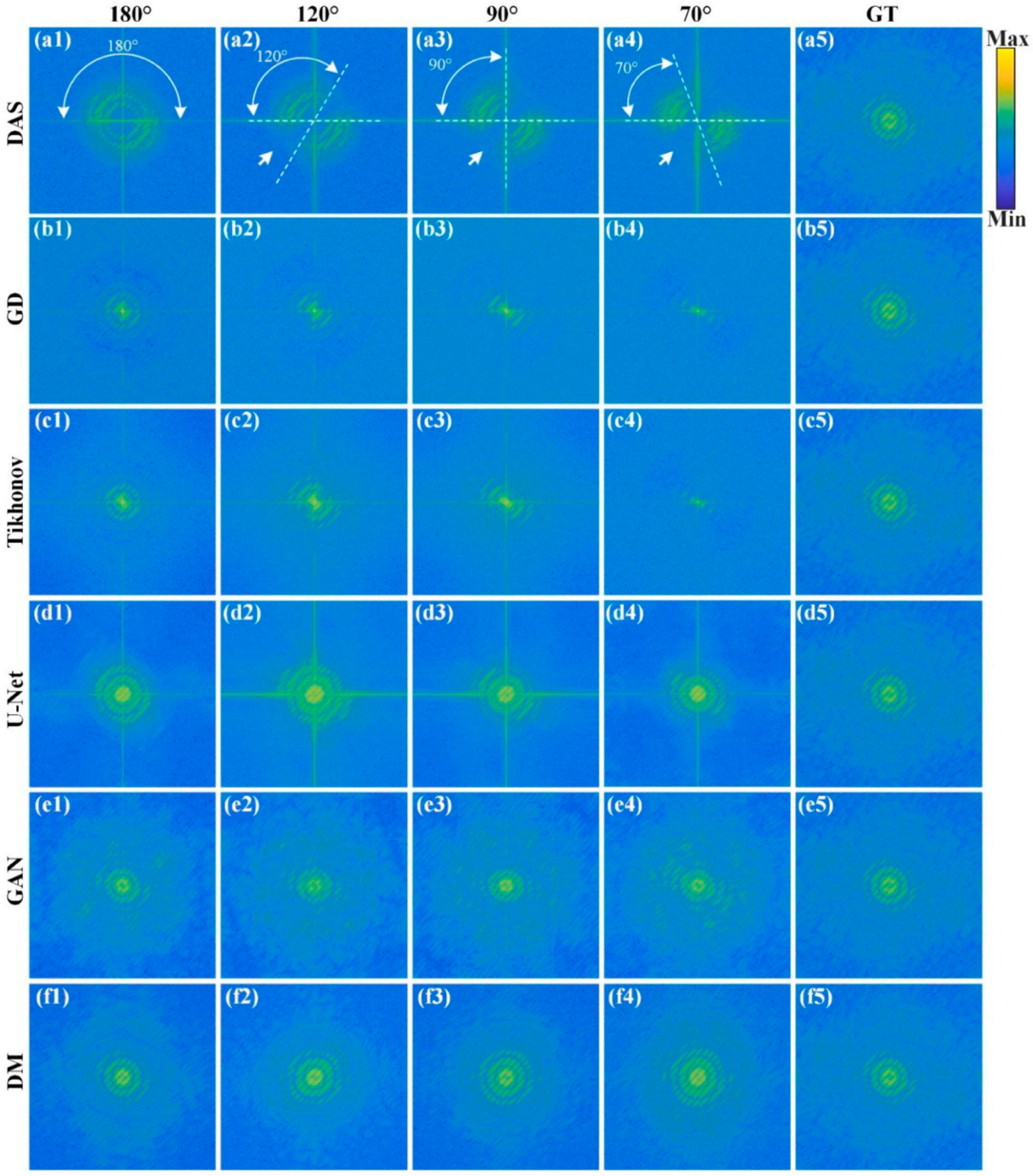


Fig. 9. The Fourier spectra of the results through Fourier transform. (a1)-(a4) depict the Fourier spectra corresponding to the reconstruction results of the DAS method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (b1)-(b4) represent the Fourier spectra corresponding to the reconstruction results of the GD method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (c1)-(c4) are the Fourier spectra corresponding to the reconstruction results of the Tikhonov method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (d1)-(d4) represent the Fourier spectra corresponding to the reconstruction results of the U-Net method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (e1)-(e4) show the Fourier spectra corresponding to the reconstruction results of the GAN method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (f1)-(f4) illustrate the Fourier spectra corresponding to the reconstruction results of the proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (a5)-(f5) are the same Fourier spectra corresponding to the same ground truth. DM, diffusion model. GD, gradient descent. GT, ground truth.

evaluated using circular phantom and *in vivo* experimental data and was compared with the U-Net and DAS methods. The experimental results demonstrate that the proposed method can effectively remove artifact and accurately restore the detailed information of the PAT images in different limited-view cases. Specifically, the proposed method surpasses the U-Net method even in exceedingly constrained limited-view cases (e.g., 70° -view). In limited-view case of 70°, the proposed method attains the PSNR of 31.57 dB and the SSIM of 0.95 for the circular phantom, exhibiting improvements of 21.16 dB (~203%) and 0.31 (~48%) over the DAS method, respectively. For *in vivo* experimental data, in limited-view case of 70°, the proposed method presents the PSNR of 27.12 dB and the SSIM of 0.76, with an improvement of 12.12 dB (~81%) and 0.30 (~65%) compared with the DAS method, respectively. This study illustrates the splendid performance of the

Table 1
The mean and standard deviation for the quantitative metrics of circular phantom data.

Limited-view	Method	PSNR (dB)		SSIM		FSIM		MSE	
		mean	standard	mean	standard	mean	standard	mean	standard
70°	DAS	10.95	0.4268	0.63	0.0343	0.72	0.0022	0.0794	0.0086
	U-Net	26.85	1.0029	0.91	0.0196	0.86	0.0162	0.0157	0.0045
	DM	31.33	1.1761	0.94	0.0122	0.87	0.0377	0.0008	0.0002
90°	DAS	11.97	0.6656	0.67	0.0279	0.72	0.0053	0.0633	0.0112
	U-Net	28.03	0.8413	0.91	0.0253	0.86	0.0128	0.0120	0.0057
	DM	33.40	0.6945	0.95	0.0112	0.89	0.0309	0.0004	0.0001
120°	DAS	12.39	0.9885	0.67	0.0338	0.73	0.0048	0.0585	0.0160
	U-Net	28.50	0.5123	0.91	0.0279	0.86	0.0176	0.0112	0.0037
	DM	33.60	1.5296	0.96	0.0130	0.91	0.0302	0.0004	0.0002
180°	DAS	14.18	0.7457	0.69	0.0306	0.74	0.0082	0.0396	0.0081
	U-Net	28.61	0.2248	0.91	0.0280	0.87	0.0130	0.0100	0.0036
	DM	33.72	1.0390	0.96	0.0080	0.94	0.0076	0.0004	0.0001

proposed method in the limited-view PAT, further extending the application scope of PAT in the biomedical field.

The proposed method can significantly accelerate the iterative reconstruction speed while maintaining high reconstruction quality. In the traditional model-based methods, the forward operator A and the adjoint operator A^* require to be computed at each iteration by utilizing k-Wave toolbox. The calculation of two operators takes lots of time [55,56]. In the current dataset, the GD method takes ~ 5 s per iteration for reconstruction. After 900 iterations, the output PSNR/SSIM of the GD method reaches 19.00 dB/0.22, with a reconstruction time of ~ 4500 s (~75 min). In the reconstruction process of the proposed method, each iteration comprises two steps: the data generation based on reverse SDE (including PC sampling) and the computation of gradient descent. The time required for one iteration is ~ 6 s. As shown in Figs. 5(d) and 5(e), the SSIM and PSNR subsequently stabilize at about the 700th iteration. Therefore, it takes ~ 4200 s (~70 min) to stabilize using the proposed method. Considering the outperformance of the proposed method, measures can be taken to accelerate the reconstruction process. On the one hand, utilizing better-performing graphics cards (e.g., NVIDIA GeForce RTX 4090) can further shorten the processing time. On the other hand, considering that the PSNR of the proposed method at the 30th iteration surpasses that of the GD method at the 900th iteration, the PSNR reaches 24.01 dB and the structural information of the image has been reconstructed quite satisfactorily at the 100th iteration. If the images are output at the 100th iteration, the reconstruction time of the proposed method would be significantly shortened to ~ 10 minutes, making it ~ 7.5 times faster than the GD method.

The diversity and scale of the dataset can impact the generalization of the proposed method. In this study, the training set was obtained from photoacoustic images reconstructed by DAS method under 512 projections (approximate to full-view), which still exhibit a few artifacts, as shown in Fig. 6(a5)-6(f5). Consequently, the photoacoustic images with slight artifacts as the gold standard images is feasible. However, when trained with images with significant artifacts (e.g., limited-view case of 180°), the reconstructed images would also have more artifacts. Therefore, to achieve superior reconstruction performance, it is suggested to use artifact-free PAT reconstruction images for training. Furthermore, all experimental datasets in this study originate from the same PAT system and the same kinds of imaging subjects. Hence, it is imperative to further enhance and validate the generalization of the proposed method, particularly when dealing with limited-view data from different PAT systems or diverse kinds of imaging subjects. By augmenting the datasets, the generalization of the proposed method can be improved. The augmented datasets should include full-view images from diverse samples obtained from different PAT systems. Due to the lack of original limited-view data from real systems, the limited-view data is obtained via the virtual PAT. This process results in lower image quality using the U-Net method of the *in vivo* data results compared to those reported in [25]. For limited-view data obtained from real systems, the forward operator A and the adjoint operator A^* for model-based iteration need to be determined using parameters (such as the scanning angle, number of transducers, the bandwidth of the transducer, the center frequency and scanning radius) from the real system [57,58]. Additionally, the proposed method can be extended to other array geometries and human data, contingent upon obtaining the forward operator A of the system.

The score-based diffusion model can estimate the probability distribution of given data through score network, without ample paired datasets for training. Possessing a flexible network structure capable of adapting to various types of data and application scenarios, the score-based diffusion model exhibits strong robustness and generalization performance. However, there are still areas for improvement in this study. Regarding the reconstruction speed, the iterative reconstruction starts from Gaussian noise and reaches stability at about the 700th iteration. In future work, the current model will be replaced by diffusion models require fewer iterations. For example, Luo *et al.* proposed an IR-SDE (Image Restoration Stochastic Differential Equation) diffusion model [59]. The iterative reconstruction starts with the degraded images and yields satisfactory results within 100 iterations. In terms of model performance, this work only utilizes a single model to learn data prior information. Yet, the prior information learned by a single model is limited, whereas multiple models offer inherent complementarity in data generation. For instance, Li *et al.* proposed a multi-modal collaborative diffusion federated learning framework called FedDiff [60]. A dual-branch diffusion model was constructed to extract data features, with two model data inputted into distinct branches of the encoder. Inspired by this, we will attempt to embed multiple models within a single reconstruction task to learn prior information from different limited-view cases and enhance model performance and reconstruction quality. Furthermore, various types of noise interference may occur in practical PAT imaging scenarios, such as the thermal acoustic noise (Gaussian noise) and the noise that originates from the ultrasonic transducers (Johnson noise) [61–63]. The noise used in the score-based diffusion model is Gaussian noise. For other noise that does not follow Gaussian distribution, it can be achieved by other kinds of diffusion models [64]. For example, Bansal *et al.* proposed a cold diffusion model [64], which can be diffused with arbitrary noise. Consequently, the cold diffusion model can be applied to a wider range of scenarios, augmenting the clinical applicability of diffusion models in PAT.

Funding

National Natural Science Foundation of China (62265011, 62122033); Jiangxi Provincial Natural Science Foundation (20224BAB212006, 20232BAB202038); National Key Research and Development Program of China (2023YFF1204302).

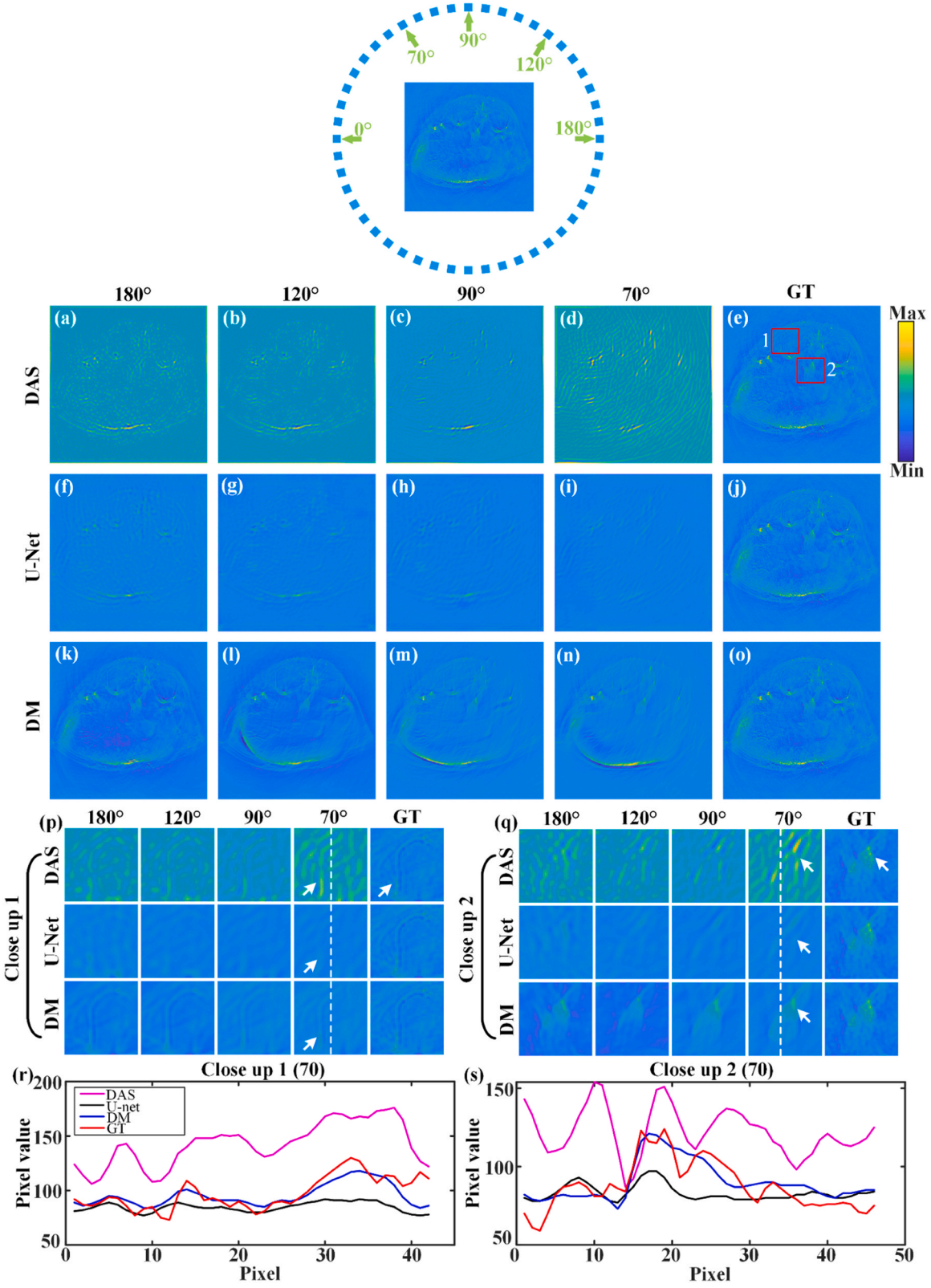


Fig. 10. The reconstruction results of the *in vivo* experimental data. (a)-(d) represent the results of the DAS method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (f)-(i) are the results of the U-Net method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (k)-(n) are the results of the proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (e), (j) and (o) are the same ground truth. (p) and (q) represent the close-up images indicated by red rectangles 1 and 2. (r)-(s) show the intensity profiles along the dashed lines in (p) and (q).

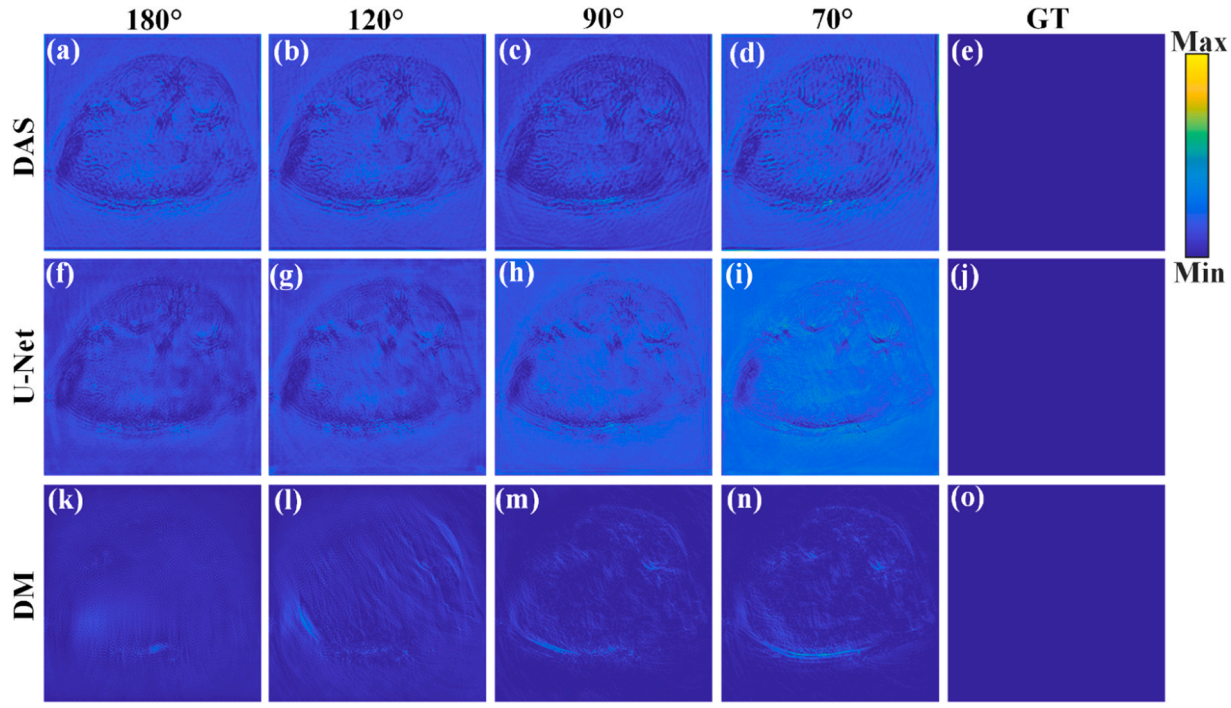


Fig. 11. (a)-(o) are the error maps of the reconstruction results. (a)-(d) depict the error maps corresponding to the reconstruction results of the DAS method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (f)-(i) represent the error maps corresponding to the reconstruction results of the U-Net method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (k)-(n) illustrate the error maps corresponding to the reconstruction results of the proposed method in limited-view cases of 180°, 120°, 90° and 70°, respectively. (e), (j) and (o) are the error maps corresponding to the same ground truth.

Table 2
The quantitative metrics for the reconstruction results of *in vivo* experimental data.

Limited-view	Method	PSNR (dB)	SSIM	FSIM	MSE
70°	DAS	15.00	0.46	0.70	0.0342
	U-Net	26.32	0.68	0.72	0.0187
	DM	27.12	0.76	0.80	0.0022
90°	DAS	17.81	0.54	0.74	0.0182
	U-Net	26.65	0.68	0.72	0.0051
	DM	29.18	0.80	0.83	0.0014
120°	DAS	16.43	0.56	0.75	0.0247
	U-Net	27.13	0.71	0.77	0.0047
	DM	27.19	0.83	0.87	0.0017
180°	DAS	16.04	0.57	0.74	0.0269
	U-Net	27.33	0.76	0.81	0.0033
	DM	28.55	0.91	0.93	0.0011

CRediT authorship contribution statement

Jiahong Li: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jiabin Lin:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Qiegen Liu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Yubin Cao:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yiguang Wang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Zhiyuan Zheng:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Wenhua Zhong:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xianlin Song:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Kangjun Guo:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Zilong Li:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Data curation, Conceptualization. **Guijun Wang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Our code is publicly available at http://github.com/yqx7150/Limited-view-PAT-Diffusion.

Acknowledgements

The authors thanks Zehao Sun from Ji luan Academy, Nanchang University for helpful discussions.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.pacs.2024.100623.

References

[1] P. Zhang, L. Li, L. Lin, P. Hu, J. Shi, Y. He, L. Zhu, Y. Zhou, V. Wang, High-resolution deep functional imaging of the whole mouse brain by photoacoustic computed tomography in vivo,", J. Biophotonics 11 (1) (2017) e201700024.

[2] L.V. Wang, J. Yao, A practical guide to photoacoustic tomography in the life sciences, Nat. Methods 13 (8) (2016) 1548–7105.

[3] S. Hu, K. Maslov, V. Tsytsarev, L.V. Wang, Functional transcranial brain imaging by optical resolution photoacoustic microscopy, J. Biomed. Opt. 14 (4) (2009) 040503.

[4] Y. Wu, J. Kang, W.G. Lesniak, A. Lisok, H.K. Zhang, R.H. Taylor, M.G. Pomper, E. M. Boctor, "System-level optimization in spectroscopic photoacoustic imaging of prostate cancer, Photoacoustics 27 (2022) 2213–5979.

[5] R. Cao, C. Zhang, V.V. Mitkin, M.F. Lankford, J. Li, Z. Zuo, J.R. Stone, S. Hu, Comprehensive characterization of cerebrovascular dysfunction in blast traumatic brain injury using photoacoustic microscopy, " J. Neurotraum 36 (10) (2019) 0897–7151.

[6] W. Zheng, D. Lee, J. Xia, Photoacoustic tomography of fingerprint and underlying vasculature for improved biometric identification, Sci. Rep. 11 (1) (2021) 17536.

[7] M.A. Lediju Bell, Photoacoustic imaging for surgical guidance: Principles, applications, and outlook, J. Appl. Phys. 128 (6) (2020) 0021–8979.

[8] M. Cui, H. Zuo, X. Wang, K. Deng, J. Luo, C. Ma, Adaptive photoacoustic computed tomography, Photoacoustics 21 (3) (2021) 2213–5979.

[9] G. Matrone, A.S. Savoia, G. Caliano, G. Magenes, The delay multiply and sum beamforming algorithm in ultrasound B-mode medical imaging, IEEE T. Med. Image 34 (4) (2015) 940–949.

[10] B.E. Treeby, E.Z. Zhang, B.T. Cox, Photoacoustic tomography in absorbing acoustic media using time reversal, Inverse Probl. 26 (11) (2010) 0266–5611.

[11] M.H. Xu, L.V. Wang, "Pulsed-microwave-induced thermoacoustic tomography: Filtered back-projection in a circular measurement configuration, Med. Phys. 29 (8) (2002) 1661–1669.

[12] Y. Xu, L.V. Wang, G. Ambartsoumian, P. Kuchment, Reconstructions in limited-view thermoacoustic tomography, Med. Phys. 31 (4) (2004) 724–733.

[13] G. Paltauf, R. Nuster, R. Burgholzer, Weight factors for limited angle photoacoustic tomography, Phys. Med. Biol. 54 (11) (2009) 0031–9155.

[14] J. Zhang, H. Lan, C. Yang, S. Guo, F. Gao, Limited-view photoacoustic imaging reconstruction with dual domain inputs under mutual information constraint,", arXiv, arXiv (2020) 2011–6147.

[15] J. Schwab, S. Antholzer, M. Haltmeier, Learned backprojection for sparse and limited view photoacoustic tomography,", in: Proc. SPIE, **10878**, 2019.

[16] G. Paltauf, J.A. Viator, S.A. Prahl, S.L. Jacques, Iterative reconstruction algorithm for optoacoustic imaging, J. Acoust. Soc. Am. 112 (4) (2002) 1536–1544.

[17] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, S. Arridge, Model-based learning for accelerated, limited-view 3-D photoacoustic tomography, IEEE T. Med. Imag. 37 (6) (2018) 1382–1393.

[18] X.L. Dean-Ben, A. Buehler, V. Ntziachristos, D. Razansky, "Accurate Model-Based Reconstruction Algorithm for Three-Dimensional Optoacoustic Tomography,", IEEE T. Med. Imaging 31 (10) (2012) 1922–1928.

[19] C. Huang, K. Wang, L. Nie, L.V. Wang, M.A. Anastasio, Full-Wave Iterative Image Reconstruction in Photoacoustic Tomography with Acoustically Inhomogeneous Media, IEEE T. Med. Imaging 32 (6) (2013) 1097–1110.

[20] X.L. Deán-Ben, V. Ntziachristos, D. Razansky, Statistical optoacoustic image reconstruction using a-priori knowledge on the location of acoustic distortions, Appl. Phys. Lett. 98 (17) (2011) 0003–6951.

[21] W. Rawat, Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review, Neural Comput. 29 (9) (2017) 2352–2449.

[22] S. Guan, A.A. Khan, S. Sikdar, P.V. Chitnis, Limited-View and Sparse Photoacoustic Tomography for Neuroimaging with Deep Learning, Sci. Rep. 10 (1) (2020) 2045–2322.

[23] T. Tong, W. Huang, K. Wang, Z. He, L. Yin, X. Yang, S. Zhang, J. Tian, Domain transform network for photoacoustic tomography from limited-view and sparsely sampled data, Photoacoustics 19 (100190) (2020) 2213–5979.

[24] H. Shahid, A. Khalid, X. Liu, M. Irfan, D. Ta, A deep learning approach for the photoacoustic tomography recovery from undersampled measurements, 1662-453X, Front. Neurosci. -Switz. 15 (2021), 1662-453X.

[25] N. Davoudi, X.L. Deán-Ben, D. Razansky, Deep learning optoacoustic tomography with sparse data, Nat. Mach. Intell. 1 (10) (2019) 453–460.

[26] D. Rezende, S. Mohamed, "Variational inference with normalizing flows, Int. Conf. Mach. Learn. (2015) 1530–1538.

[27] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courvile, Y. Bengio, Generative adversarial networks, Commun. ACM 63 (11) (2020) 139–144.

[28] J. Ho, A. Jain, P. Abbeel, "Denoising diffusion probabilistic models, Adv. Neural Inf. Process. Syst. 33 (2020) 6840–6851.

[29] T. Lu, T. Chen, F. Gao, B. Sun, V. Ntziachristos, J. Li, "LV-GAN: A deep learning approach for limited-view optoacoustic imaging based on hybrid datasets, 2020-0325, J. Biophotonics 14 (2) (2020), 2020-0325.

[30] M. Lucic, K. Kurach, M. Michalski, S. Gelly, O. Bousquet, "Are gans created equal? a large-scale study,", Adv. Condens. Matter Phys. 31 (2018).

[31] Y. Song, J. Sohl-Dickstein, D.P. Kingma, A. Kumar, "Score-based generative modeling through stochastic differential equations,", arXiv, arXiv (2020) 2011–13456.

[32] P. Dhariwal, A. Nichol, "Diffusion models beat gans on image synthesis, Adv. Neural Inf. Process. Syst. 34 (2021) 8780–8794.

[33] H. Chung, J.C. Ye, Score-based diffusion models for accelerated MRI, Med. Image Anal. 80 (102479) (2022) 1361–8415.

[34] X. Song, G. Wang, W. Zhong, K. Guo, Z. Li, X. Liu, J. Dong, Q. Liu, "Sparse-view reconstruction for photoacoustic tomography combining diffusion model with model-based iteration, Photoacoustics 33 (2023) 2213–5979.

[35] K. Guo, Z. Zheng, G. Wang, X. Song, "Generative model for limited-view photoacoustic tomography reconstruction, Proc. SPIE 12972 (2024) 66–72.

[36] A. Oraevsky, R. Su, H. Nguyen, J. Moore, Y. Lou, S. Bhadra, L. Forte, M. Anastasio, W. Yang, "Full-view 3D imaging system for functional and anatomical screening of the breast, Proc. SPIE 10494 (2018) 217–226.

[37] S. Na, J.J. Russin, L. Lin, X. Yuan, P. Hu, K.B. Jann, L. Yan, K. Maslov, J. Shi, D. J. Wang, C.Y. Liu, L.V. Wang, Massively parallel functional photoacoustic computed tomography of the human brain, Nat. Biomed. Eng. 6 (5) (2022) 584–592.

[38] F. Dreier, S. Pereverzyev Jr, M. Haltmeier, Operator learning approach for the limited view problem in photoacoustic tomography, Comput. Methods Appl. Math. 19 (4) (2019) 749–764.

[39] T. Wang, M. He, K. Shen, W. Liu, C. Tian, Learned regularization for image reconstruction in sparse-view photoacoustic tomography, Biomed. Opt. Express 13 (11) (2022) 5721–5737.

[40] J. Gröhl, M. Schellenberg, K. Dreher, L. Maier-Hein, Deep learning for biomedical photoacoustic imaging: A review, Photoacoustics 22 (2021) 2213–5979.

[41] S. Antholzer, M. Haltmeier, Discretization of Learned NETT Regularization for Solving Inverse Problems, 2313-433X, J. Imaging 7 (11) (2021), 2313-433X.

[42] J. Wang, Y.Y. Wang, Photoacoustic imaging reconstruction using combined nonlocal patch and total-variation regularization for straight-line scanning, 1475-925X, Biomed. Eng. Online 17 (1) (2018), 1475-925X.

[43] M.E. Hochstenbach, L. Reichel, Fractional Tikhonov regularization for linear discrete ill-posed problems, BIT Numer. Math. 51 (1) (2011) 1572–9125.

[44] O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in, Med. Image Comput. Comput. -Assist. Interv. –MICCAI 2015: 18th Int. Conf. (2015) 234–241.

[45] Z. Niu, G. Zhong, H. Yu, "A review on the attention mechanism of deep learning, Neurocomputing 452 (2021) 48–62.

[46] C. Saharia, W. Chan, H. Chang, C. Lee, J. Ho, T. Salimans, D. Fleet, M. Norouzi, "Palette: Image-to-image diffusion models, ACM SIGGRAPH 2022 Conf. Proc. (2022) 1–10.

[47] A. Brock, J. Donahue, K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," arXiv, arXiv:1809.11096 (2018).

[48] X. Song, W. Zhong, Z. Li, S. Peng, H. Zhang, G. Wang, J. Dong, X. Liu, X. Xu, Q. Liu, Accelerated model-based iterative reconstruction strategy for sparse-view photoacoustic tomography aided by multi-channel autoencoder priors, " J. Biophotonics 17 (1) (2024) e202300281.

[49] B. Huang, L. Zhang, S. Lu, B. Lin, W. Wu, Q. Liu, "One sample diffusion model in projection domain for low-dose CT imaging," arXiv, arXiv:2212.03630 (2022).

[50] C. Matthews, J. Weare, "Langevin Markov Chain Monte Carlo with stochastic gradients," arXiv, arXiv:1805.08863 (2018).

[51] B.E. Treeby, B.T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, 021314-021314, J. Biomed. Opt. 15 (2) (2010), 021314-021314.

[52] J.Y. Zhu, T. Park, P. Isola, A.A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in 2017, IEEE Int. Conf. Comput. Vis. (ICCV) (2017) 2223–2232.

[53] M. Lu, X. Liu, C. Liu, C. Liu, B. Li, W. Gu, J. Jiang, D. Ta, Artifact removal in photoacoustic tomography with an unsupervised method, Biomed. Opt. Express 12 (10) (2021) 6284–6299.

[54] L. Zhang, L. Zhang, X. Mou, D. Zhang, "FSIM: A feature similarity index for image quality assessment,", IEEE Trans. Image Process. 20 (8) (2011) 2378–2386.

[55] A. Hauptmann, B. Cox, "Deep learning in photoacoustic tomography: current approaches and future directions, 112903-112903, J. Biomed. Opt. 25 (11) (2020), 112903-112903.

[56] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, S. Arridge, Model-based learning for accelerated, sparse-view 3-D photoacoustic tomography, IEEE Trans. Med. Imag. 37 (6) (2018) 1382–1393.

[57] J. Adler, O. Öktem, Solving ill-posed inverse problems using iterative deep neural networks, Inverse Prob 33 (12) (2017) 124007.

[58] A. Hauptmann, B. Cox, F. Lucka, N. Huynh, M. Betcke, P. Beard, S. Arridge, Approximate k-space models and deep learning for fast photoacoustic reconstruction, Lect. Notes Comput. Sci. 11074 (2018) 103–111.

[59] Z. Luo, F.K. Gustafsson, Z. Zhao, J. Sjölund, T.B. Schön, "Image restoration with mean-reverting stochastic differential equations,", 2301.11699, arXiv, arXiv (2023), 2301.11699.

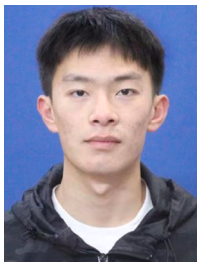
[60] D.X. Li, W. Xie, Z.X. Wang, Y.B. Lu, Y.S. Li, L.Y. Fang, "FedDiff: Diffusion Model Driven Federated Learning for Multi-Modal and Multi-Clients," arXiv, arXiv: 2401.02433 (2023).

[61] G. Guney, N. Uluc, A. Demirkiran, E. Aytac-Kipergil, M.B. Unlu, O. Birgul, "Comparison of noise reduction methods in photoacoustic microscopy, Comput. Biol. Med. 109 (2019) 333–341.

[62] E.R. Hill, W. Xia, M.J. Clarkson, A.E. Desjardins, Identification and removal of laser-induced noise in photoacoustic imaging using singular value decomposition, Biomed. Opt. Express 8 (1) (2017) 68–77.

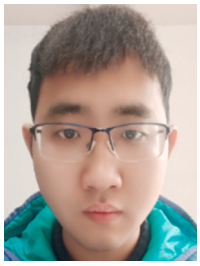
[63] A.M. Winkler, K.I. Maslov, L.V. Wang, Noise-equivalent sensitivity of photoacoustics, J. Biomed. Opt. 18 (9) (2013) 097003.

[64] A. Bansal, E. Borgnia, H.M. Chu, J.S. Li, H. Kazemi, F. Huang, M. Goldblum, J. Geiping, T. Goldstein, "cold diffusion: Inverting arbitrary image transforms without noise"



Kangjun Guo is currently studying for bachelor degree in Automation Science in Nanchang University, Nanchang, China. His research interests include image processing, artificial intelligence and photoacoustic tomography.

Zhiyuan Zheng received the bachelor degree in Electronic Information Engineering from Nanchang University, Nanchang, China. He is currently studying in Nanchang University for master's degree in Information and Communication Engineering. His research interests include deep learning, image processing and photoacoustic tomography.



Wenhua Zhong received the bachelor degree in Electronic Information Engineering from Jiangxi Science and Technology Normal University, Nanchang, China. He is currently studying in Nanchang University for master's degree in Electronic Information. His research interests include computer vision, deep learning, photoacoustic microscopy and photoacoustic tomography.

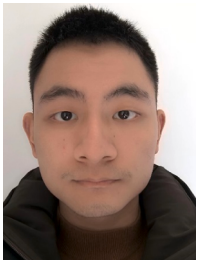
Zilong Li received the bachelor degree in Electronic Information Engineering from Guilin University of Electronic Technology, Guilin, China. He is currently studying in Nanchang University for master's degree in Electronic Information Engineering. His research interests include optical imaging, deep learning and photoacoustic tomography.

Guijun Wang received the bachelor degree in Electronic Information Science and Technology from Nanchang Hangkong University, Nanchang, China. He is currently studying in Nanchang University for master's degree in Electronic Information Engineering. His research interests include image processing, deep learning and photoacoustic tomography.

Jiahong Li received the bachelor degree in Electronic Information Engineering from Yanshan University, Qinhuangdao, China. He is currently studying in Nanchang University for master's degree in Electronic Information Engineering. His research interests include optical imaging, deep learning and photoacoustic tomography.

Yubin Cao received the bachelor degree in Biomedical Engineering from Gannan Medical University, Ganzhou, China. He is currently studying in Nanchang University for master's degree in Biomedical Engineering. His research interests include optical imaging, deep learning and photoacoustic microscopy.

Yiguang Wang received the bachelor degree in Biomedical Engineering from Gannan Medical University, Ganzhou, China. He is currently studying in Nanchang University for master's degree in Biomedical Engineering. His research interests include image processing, deep learning and photoacoustic microscopy.



Jiabin Lin is currently studying for bachelor degree in Automation Science in Nanchang University, Nanchang, China. His research interests include optical imaging, deep learning and photoacoustic tomography.

Qiegen Liu received his PhD degree in Biomedical Engineering from Shanghai Jiao Tong University, Shanghai, China in 2012. Currently, he is a professor at Nanchang University. He is the winner of Excellent Young Scientists Fund. He has published more than 50 publications and serves as committee member of several international and domestic academic organizations. His research interests include artificial intelligence, computational imaging and image processing.

Xianlin Song received his PhD degree in Optical Engineering from Huazhong University of Science and Technology, China in 2019. He joined School of Information Engineering, Nanchang University as an assistant professor in Nov. 2019. He has published more than 20 publications and given more than 15 invited presentations at international conferences. His research topics include optical imaging, biomedical imaging and photoacoustic imaging.