Research article

Image reconstruction of multispectral sparse sampling photoacoustic tomography based on deep algorithm unrolling

Jia Ge [a,b,c,1], Zongxin Mo [a,b,c,1], Shuangyang Zhang [a,b,c], Xiaoming Zhang [a,b,c], Yutian Zhong [a,b,c], Zhaoyong Liang [a,b,c], Chaobin Hu [a,b,c], Wufan Chen [a,b,c,*], Li Qi [a,b,c,*]

[a] *School of Biomedical Engineering, Southern Medical University, 1023 Shatai Rd., Baiyun District, Guangzhou, Guangdong 510515, China*
[b] *Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, 1023 Shatai Rd., Baiyun District, Guangzhou, Guangdong 510515, China*
[c] *Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Southern Medical University, 1023 Shatai Rd., Baiyun District, Guangzhou, Guangdong 510515, China*

ARTICLE INFO

ABSTRACT

Photoacoustic tomography (PAT), as a novel medical imaging technology, provides structural, functional, and metabolism information of biological tissue *in vivo*. Sparse Sampling PAT, or SS-PAT, generates images with a smaller number of detectors, yet its image reconstruction is inherently ill-posed. Model-based methods are the state-of-the-art method for SS-PAT image reconstruction, but they require design of complex handcrafted prior. Owing to their ability to derive robust prior from labeled datasets, deep-learning-based methods have achieved great success in solving inverse problems, yet their interpretability is poor. Herein, we propose a novel SS-PAT image reconstruction method based on deep algorithm unrolling (DAU), which integrates the advantages of model-based and deep-learning-based methods. We firstly provide a thorough analysis of DAU for PAT reconstruction. Then, in order to incorporate the structural prior constraint, we propose a nested DAU framework based on plug-and-play Alternating Direction Method of Multipliers (PnP-ADMM) to deal with the sparse sampling problem. Experimental results on numerical simulation, *in vivo* animal imaging, and multispectral unmixing demonstrate that the proposed DAU image reconstruction framework outperforms state-of-the-art model-based and deep-learning-based methods.

1. Introduction

Multispectral photoacoustic tomography (PAT) is a hybrid imaging method [1,2] that has demonstrated its capability in visualizing mammalian brain neuronal activity [3], tracking tumor advancement [4,5], diagnosing issues related to placental and fetal health [6,7]. However, in order to obtain multi-wavelength photoacoustic data, the laser need to continuously switch excitation wavelengths for repeated scanning, which places a high demand on multi-channel simultaneous acquisition and the system cost [2]. To alleviate this, the sparse sampling strategy can be used, which reduces the total number of detectors. However, the reduction of detectors leads to significant degradation in the quality of reconstructed images. The popular universal back-projection (UBP) algorithm [8] has been shown to perform poorly in sparse sampling PAT imaging. Therefore, specific image reconstruction algorithms are needed to improve the quality of sparse sampling PAT [9–11]. Current image reconstruction methods for sparse sampling PAT include the model-based (MB) methods and deep-learning-based (DL) methods.

Model-based methods treat the PAT image reconstruction problem as a linear inverse problem, which usually contains a data fidelity term and a regularization term [12,13]. Prior knowledge of the image can be incorporated though the regularization term to address the sparse sampling problem [14]. For example, total variational (TV) regularization combined with L1-norm and image difference can improve PAT image reconstruction from sparse sampled data [15,16]. Non-local means (NLM) regularization enables enhanced reconstruction of PAT images by searching for similar image patches and performing weighted

---

summation [2,17]. Furthermore, by incorporating the edge information of the structural prior into the sparse sampling PAT images, the directional total variation (dTV) method avoids traversing of the image and obtains promising results with reduced computational cost [13]. Nonetheless, all the aforementioned approaches require the design of analytical and intricate handcrafted regularization terms, and inappropriate prior selection can lead to unsatisfactory performance or even reconstruction failure.

In recent years, deep learning has gained prominence in both industrial and medical imaging applications [18–20]. For PAT image reconstruction, Waibel et al. [21] first proposes to utilize the U-Net framework to realize direct image reconstruction, which harnesses the time series data to obtain initial pressure distribution without pre-processing. Lan et al. [22] proposes to employ the general adversarial network for reconstruction, where raw data reconstructed by conventional reconstruction algorithm is also fed into the network to leverage additional information. However, the performance of learning-based methods is crucially relying on well-labeled real-world training data, which is typically deficient in PAT imaging. Furthermore, DL network model is often regarded as a black box, which lacks interpretability.

The above problems can be well addressed by deep algorithm unrolling (DAU), a novel DL-based method that combines the advantages of both the model-based and the DL-based method [23,24]. The main idea of DAU is that conventional model-based iterative optimization algorithm can be implemented by convolutional neural network (CNN) blocks and deploys learnable parameters in each iteration, so that the parameters could be updated in an end-to-end manner [25–27]. For instance, Schwab et al. [28] incorporates supplementary weight factors that considers the properties of acoustic detectors into the standard back-projection approach to improve reconstruction quality. Hauptmann et al. [29] introduces an iterative learning strategy that involves iterative reconstruction by training distinct networks for each iteration step and incorporating them into the reconstruction process. Although these methods have demonstrated the feasibility of DAU for PAT image reconstruction, they have not employed the rich prior information provided by multispectral excitation in multispectral PAT.

To address the above limitation, here we propose a new DAU-based image reconstruction method that incorporates the structural prior into the unrolling framework for multispectral sparse sampling PAT. Our DAU framework is based on the unrolling of alternating direction method of multipliers (ADMM) [30]. We choose the dTV constraint as the tool for introducing structural information. We transform the dTV prior to an image restoration block, which is then integrated into the unrolled ADMM framework in a plug-and-play (PnP) manner. This eventually leads to a DAU network model nested by dTV, or NDU-dTV as we named it. To validate the effectiveness of our method, we first test our NDU-dTV model in numerical simulation experiment with comparison to other model-based and DL-based methods under different sparse sampling rates. Further, we perform multispectral PAT imaging experiment on mice *in vivo*, where our NDU-dTV method shows significant improvement in both reconstruction results and spectral un-mixing accuracy, thus revealing its attractive potential.

2. Theory

2.1. PAT imaging model

In model-based PAT image reconstruction, the relationship between the photoacoustic signal and the image can be represented in the form of a discrete matrix, defined as [12]

$$p = Ax, \tag{1}$$

where p and x are the original photoacoustic signal and the target image represented by vectors. A is the system matrix, with its elements rep-

resenting the response function mapping from object space to data space. The solution of x can be obtained by solving the model-based optimization problem:

$$x = arg\min \frac{1}{2}\|p - Ax\|_2^2 + \lambda J(x), \tag{2}$$

where $\|p - Ax\|$ is the data fidelity term; $J(x)$ is a regularization term and λ is the trade-off parameter balancing the data fidelity term and regularization term.

2.2. Multispectral interlaced sparse sampling PAT method

Current multispectral PAT system employs transducer array encompassing hundreds of detection elements in order to ensure effective angular coverage and facilitates multispectral measurement [31]. In practice, the reduction of data volume could be achieved by sparse sampling, but it is challenging to retain detail structural information when there is only a very small number of detectors available. In multispectral imaging, the missing structural information problem can be addressed through the technique called interlaced sparse sampling photoacoustic tomography (ISS-PAT) [2]. In ISS-PAT, as shown in Fig. 1, the sparse detector arrays rotate by a certain angle every time the excitation wavelength is switched from one another. Following a certain number of rotations, the obtained sparse signals contain information from different scanning angles and different wavelengths. The signals are then combined and reconstructed into a structural prior image, which is then used to guide the reconstruction of images at individual wavelength. Compared to dense sampling, the ISS mode reduces the
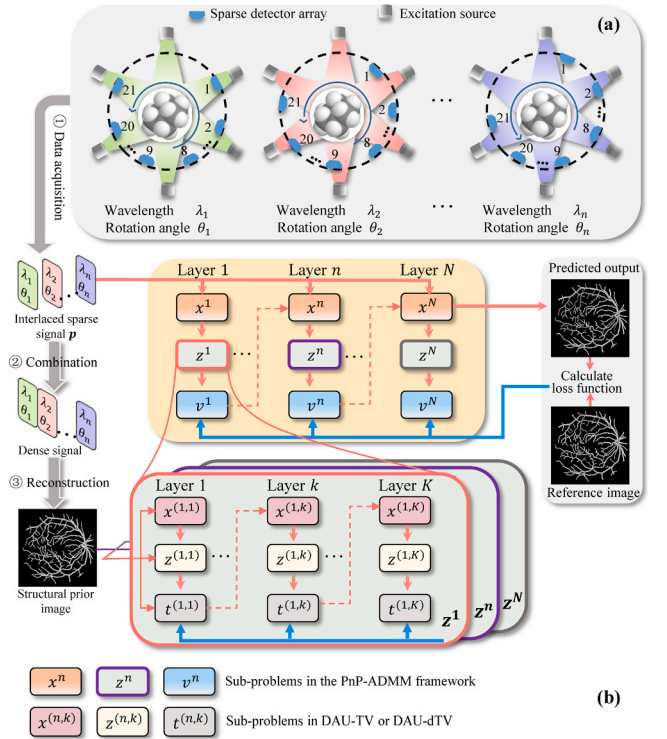

Fig. 1. (a) Schematic diagram of multispectral ISS-PAT principle. ① Acquiring interlaced sparse-sampling signals, where the detector array rotates with wavelength switching. ②Linearly combining the acquired sparse signals to obtain a set of dense signal. ③ The dense signal is reconstructed to obtain a structural prior image. (b) Overview of the proposed NDU-dTV for SS-PAT. The top orange-colored box denotes the unrolled ADMM network for image reconstruction. Underneath is the dTV-based image restoration blocks nested into the unrolled ADMM, which incorporates the prior information. K and N represent the number of the stages in image restoration block and the iterations in ADMM respectively.

detector number but does not increase the total multispectral imaging time.

In ISS-PAT image reconstruction, the structural prior is introduced by designing a regularization term to leverage the structural similarity between the image to be reconstructed and the structural prior image. To do so, the directional total variational (dTV) method has been proposed [13] for efficient image reconstruction. The dTV regularization assesses the need for constraint by evaluating the gradient direction of the reconstructed image and the prior image at a specific point, it can be written as

$$J_{dTV}(x, a) = |P_\xi \nabla x|, \tag{3}$$

where $P_\xi \nabla x$ is used to extract the structural information from the prior image, and defined as

$$P_\xi \nabla x = P_\xi Dx = \nabla x - \langle \xi, \nabla x \rangle \xi, \tag{4}$$

∇x is the gradient direction of the image to be reconstructed, and ξ is the unit gradient direction of the structural prior image, defined as

$$\xi = \frac{\nabla a}{\sqrt{|\nabla a|^2 + \varepsilon^2}}, \tag{5}$$

where a is the structural prior image; ε is a constant term that prevents the denominator part from being zero. As can be seen, when the angle between ∇x and ξ is 0, then $\langle \xi, \nabla x \rangle \xi = \nabla x$ and $P_\xi \nabla x = 0$, and $J_{dTV}(x, a)$ is minimum. This means that the structure of the image to be reconstructed resembles that of the prior image. In this way, because the prior image is only used to measure the similarity of the gradient direction, it does not involve in the weighted average operation. Therefore, the intensity of the prior image will not affect the intensity of the target image [13].

3. Method

3.1. Deep algorithm unrolling based on ADMM

The model-based image reconstruction problem shown in Eq.(2) could be solved by various algorithms, such as iterative shrinkage-thresholding algorithm (ISTA) [32], alternating direction method of multipliers (ADMM) [33], and primal-dual hybrid gradient (PDHG) algorithm [34]. Among them, ADMM is a method that solves generic unconstrained optimization problems with non-differentiable convex functions. It combines the robustness of the augmented Lagrangian relaxation method with the effectiveness of the multiplier method, leading to an efficient and reliable convergence process [35]. It is the first algorithm that has been demonstrated with deep algorithm unrolling [30].

As a variant of the ADMM, Plug-and-Play ADMM (PnP-ADMM) has shown enhanced performance compared to the standard ADMM without explicit regularization in image reconstruction tasks [36,37]. PnP-ADMM exhibits faster convergence rates and requires fewer iterations compared to traditional ADMM, particularly in image processing tasks [38]. By leveraging domain-specific denoisers, PnP-ADMM can exploit the structure of the underlying problem more effectively, leading to improved convergence behavior and reduced computational complexity [39]. Introducing $z = \nabla x = Dx$, the problem of PAT reconstruction in Eq. (2) can be solved by augmented Lagrangian function as

$$L_\rho(x, z, v) = \frac{1}{2}\|p - Ax\|_2^2 + \beta|z| + \lambda^T(D_1 x - z) + \frac{\rho}{2}\|D_1 x - z\|_2^2 \tag{6}$$

The above problem can be solved by DAU based on PnP-ADMM as

$$\begin{cases} x^n = (A^T A + \rho D_1^T D_1)^{-1}(A^T p + \rho D_1^T(z^{n-1} - v^{n-1})), \\ \qquad\qquad z^n = G(x^n, v^{n-1}), \\ \qquad v^n = v^{n-1} - \sigma_1 z^n + \sigma_2 D_1 x^n, \end{cases} \tag{7}$$

where G denotes the off-the-shelf denoiser. ρ, σ_1, σ_2 are hyper-parameters.

Recent researches have shown that we can implicitly use CNN to represent the convolution operation D_l, D_l^T. As illustrated in the yellow box in Fig. 1(b), by generalizing the iterative ADMM algorithm to a deep architecture, its iterative procedures in Eq. (7) are mapped to a data flow graph, so that the $n - th$ iteration of ADMM algorithm corresponds to the $n - th$ stage of the data flow graph. This generalization process is referred to as *deep algorithm unrolling*, because it spreads the ADMM steps in the manner of deep networks. In contrast to conventional black-box-like CNN models, DAU network is considered interpretable because it is stemmed from rigorous optimization algorithm. For PnP-ADMM, the denoiser G can also be unrolled into the DAU network as an end-to-end trainable deep network and typically yields superior results with fewer iterations [40].

3.2. PAT image reconstruction based on NDU-dTV

In our SS-PAT image reconstruction task, although the dTV method could achieve comparable performance, it is non-trivial to tune the iterative parameters in ADMM solver. Also, the iteration involves complex iterative computation scheme and specific convergence condition. Introducing the dTV method by using PnP-ADMM may ease these limitations.

Assuming the output of the reconstruction layer as y, the denoising problem can be solved as

$$\begin{cases} x^{(n, \ k)} = \left[I + \mu D_2^T P_\xi^T P_\xi D_2\right]^{-1}[y + \mu P_\xi D_2(z^{(n,k-1)} - t^{(n,k-1)})], \\ z^{(n,k)} = \beta_1 z^{(n,k-1)} + \beta_2(x^{(n,k)} + t^{(n,k-1)}) - D_2^T P_\xi^T H(P_\xi D_2 z^{(n,k-1)}), \\ \qquad\qquad t^{(n,k)} = t^{(n,k-1)} - \gamma(z^{(n,k)} - P_\xi D_2 x^{(n,k)}), \end{cases} \tag{8}$$

where H represents the non-linear transforms. $\mu, \gamma, \beta_1, \beta_2$ are the learnable parameters.

Next, as stated above, the denoiser in PnP-ADMM can be implemented by a DAU network. In our case, this means we can integrate the dTV method into the deep algorithm unrolling, resulting in DAU-dTV as Eq. (8). Inspired by the recently reported nested DAU method [27], we incorporate the DAU-dTV that serves as an image restoration module with the inclusion of structural prior information, into the unrolled ADMM framework. As shown in Fig. 1(b), each stage of the network consists of a reconstruction layer (X^n), an image restoration block (Z^n) and a variable update layer (V^n), corresponding to the iteration of PnP-ADMM. Because of its nested structure, we call our model NDU-dTV.

Image reconstruction layer (X^n). This layer reconstructs an image according to the reconstruction operation X^n in Eq. (7). The data flow of this layer in $n - th$ iteration is illustrated in Fig. 2(a). The output x^n in this layer is the input for subsequent image restoration block Z^n. We do not constrain the parameters to be the same in different iterations to enhance network capacity.

Image restoration block (Z^n). We next introduce an image restoration block to implement Z^n, which is replaced by denoising method based on DAU illustrated in Eq. (8). According to the second subproblem in Eq.(8), we first decompose and generalize the operation $D^T P_\xi^T H(P_\xi Dz^{(n,k-1)})$ to the residual CNN structure. As shown in Fig. 2(b), for the $k - th$ stage in the image restoration block, the residual CNN structure is defined as 'Convolution (*Conv.*) +Batch Normalization (*BN*)
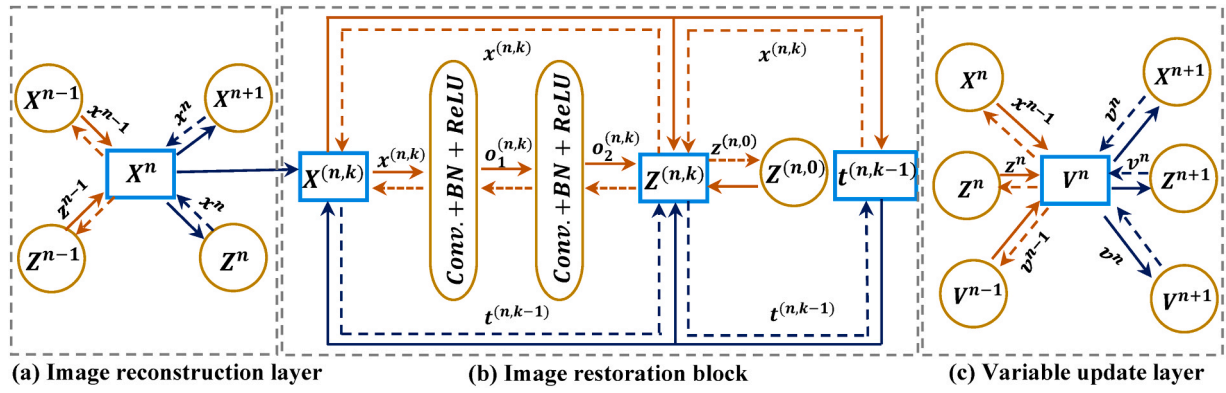
Fig. 2. The data flow of NDU-dTV framework. The rectangular box represents the focal layer, while the circles symbolize the layers interconnected with the focal layer. Solid arrows signify the data propagation during the forward pass, and dashed arrows represent the reverse update during gradient computation in back-propagation.

+Rectified Linear Unit (*ReLU*)'.

As shown in Fig. 2(b), given the outputs of previous layers $x^{(n,k)}$, $t^{(n,k-1)}, z^{(n,k-1)}$ and $o_2^{(n,k)}$, the output of this layer is $z^{(n,k)} = \beta_1 z^{(n,k-1)} + \beta_2 (t^{(n,k-1)} + t^{(n,k-1)}) - o_2^{(n,k)}$, where $o_2^{(n,k)}$ represents the output of the residual CNN structure.

Variable update layer (V^n). We generalize the variable update procedure V^n in Eq. (7) to the multiplier update layer. As shown in Fig. 2 (c), given three inputs x^n, z^n and v^{n-1}, the output of variable update layer in iteration n is defined as $v^n = v^{n-1} - \sigma_1 z^n + \sigma_2 D_1 x^n$.

4. Experimental setup

4.1. NDU-dTV implementation

The optimization of the NDU-dTV is implemented using PyTorch, and all experiments are run on a workstation equipped with a Titan X GPU. We use the Adam solver with the initial learning rate of 1×10^{-3} with a poly decay strategy. The batch size and maximum iterative number are set to 14 and 10,000. After testing different parameter settings, we set the number of iterations $N = 10$ and stage $K = 10$. For the learnable parameters, we use the random initialization to initialize the network parameters. To assess the structural constraint provided by the dTV method, we also implemented the Total Variation (TV) method, which is known for its effective denoising performance but lacks the capability of incorporating structural prior information from external sources. The TV method is integrated into the PnP-ADMM framework by

DAU using the same setting, and thus is referred to as DAU-TV.

4.2. PAT data acquisition

All the animal experiments are carried out on a commercial multi-spectral PAT system, the MSOT inVision128 (iThera Medical GmbH, Germany). Fig. 3(a) shows a diagram of the imaging chamber of the system. In this system, a tunable laser (670–960 nm) with repetition frequency of 10 Hz, pulse width < 10 ns, and a maximum pulse energy of 60 mJ at 700 nm excites the samples. The ultrasonic signals generated from the samples are transmitted by water, and subsequently captured by the ring-shaped transducer array with a center frequency of 5 MHz and 6-dB bandwidth of 100 %. Fig. 3(b) shows the schematic diagram of the ultrasound transducer setting, which is equipped with 128 elements with 270-degree detection angle and a radius of 40.5 mm. Sparse sampling data is obtained by simulating the evenly-spaced sparse detector setting at two down-sampling rates: 1/4 and 1/6, which corresponds to 32 and 21 transducer elements respectively.

4.3. Simulation setup

Fig. 4 shows the simulation process to obtain initial PAT images. The imaged object is a vascular network from the public fundus oculi vessel datasets DRIVE [41] and STARE [42], which consisted of two mediums, oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb), located in the left and right half of the image respectively. μ_a and μ_s of HbO_2 and Hb at different wavelengths are determined by following reference [43]. The
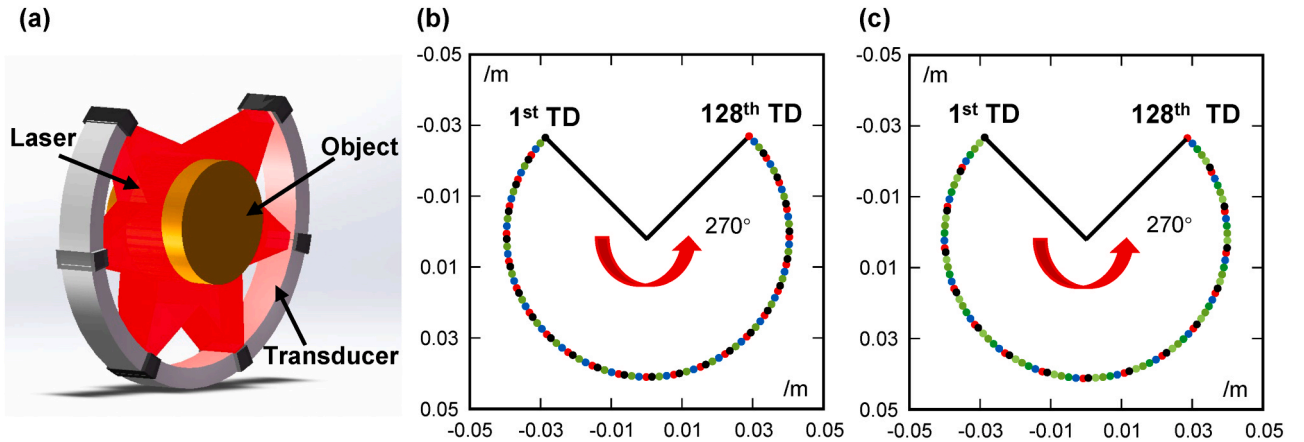

Fig. 3. (a) A side view of the cross-sectional PAT system; (b) Schematic of the array transducer setting at 1/4 sparse sampling rate, i.e. Nd=32. (c) Schematic of the array transducer setting at 1/6 sparse sampling rate, i.e. Nd=21. Dots in the same color indicate transducer elements used for imaging at an individual wavelength in the interlaced sparse sampling mode. TD: transducer.
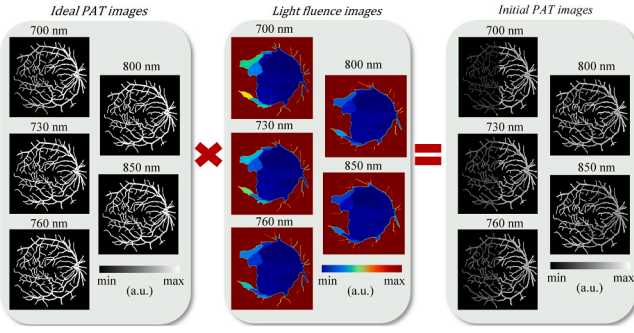
Ideal PAT images   *Light fluence images*   *Initial PAT images*

Fig. 4. The process of obtaining the initial PAT images in the simulation experiments.

reduced scattering coefficient is set as $\mu_s' = \mu_s(1 - g)$, where g is the scattering anisotropy. We set the effective attenuation coefficient u_{eff} as the absorption signals of simulation by $\mu_{eff} = \sqrt{3(\mu_\alpha + \mu_s')}$. To simulate light attenuation in the background water, we assume very low absorption and scattering, with coefficient values of $0.0001\ cm^{-1}$. These absorption and scattering values are then integrated into the Finite Element Method model using the TOAST++ toolkit [44], which subsequently calculates the fluence of each node within the entire mesh. The light source is configured according to the illumination geometry of our MSOT system, and the light fluence maps at the following five wavelengths are obtained: 700, 730, 760, 800, 850 nm. Subsequently, assuming a uniform Grüneisen parameter in soft biological tissues, we obtain the initial PAT image by multiplying the ideal PAT images with the light fluence images. Finally, we utilize the discrete forward model to produce the raw PAT signal.

Following the interlaced multispectral sparse sampling setting [13], the sparse-sampling signals are acquired at 1/4 and 1/6 sparse sampling rates. The sparse-sampling signals are set as the input for all the network models, while the initial PAT images are designated as the reference images. Additionally, the structural prior image of the dTV regularization is reconstructed using the unregularized model-based method solved by conventional gradient algorithm [45]. The dataset for the simulation experiment consists of 1550 training sets, 100 validation sets and 150 test sets.

4.4. Animal experiment

In vivo animal imaging experiment is conducted utilizing four healthy nude mice (female, 8 weeks) and dense-sampling signals are acquired from the abdomen section in the previously mentioned commercial PAT system, which is approved by Southern Medical University and follows the local Animal Ethics Committee guidelines. During the experiments, nude mice are anesthetized using a medical oxygen mixed with 3 % isoflurane (RWD Life Science Co., Ltd, China) to decrease respiratory motion artifacts. The sparse-sampling signals are set as the input for all the network models, while the reference images are reconstructed from dense-sampling signals at each wavelength using the unregularized model-based method solved by conventional gradient descent algorithm [45]. Additionally, the structural prior image of the dTV regularization is reconstructed using the dense signal linearly combined from interlaced sparse sampled at multiple wavelengths as illustrated in Fig. 1(a). The reconstruction process follows the same method as the reference images. The setting is adopted for the rest of the paper. The dataset for the animal experiment consists of 1050 training sets, 100 validation sets and 100 test sets.

4.5. Public human arms dataset

To illustrate the feasibility of the DAU in the multispectral PAT reconstruction, we measure the proposed algorithm performance on the public human arms dataset OADAT [46] and compare them with other DL and MB-based methods. The OADAT dataset is collected by multisegment array detectors which is a combination of linear array (128 elements) and concave parts on the right and left sides (64 elements), totaling to 256, from nine volunteers at five wavelengths (700, 730, 760, 800 and 850 nm) for human arms. Following the multispectral sparse sampling setting [2], the sparse-sampling signals are acquired at 1/4 and 1/6 sampling rates. The public human arms dataset consists of 400 training sets, 50 validation sets and 50 test sets. The results are presented in Supplementary Fig. S1 and Table S1.

4.6. Spectral un-mixing

For multispectral PAT imaging, the concept of spectral un-mixing is to obtain relative concentrations of light absorbers based on their spectral attributes at each wavelength. For a given wavelength λ_i, the light absorption h of the tissue at position r can be expressed as,

$$h(\lambda_i, r) = \Phi(\lambda_i, r)\mu_\alpha(\lambda_i, r) = \Phi(\lambda_i, r)\sum_{j=1}^{S}(\varepsilon_j(\lambda_i)c_j(r)) \tag{9}$$

where $\Phi(\lambda_i, r)$ is the local light fluence dependent on wavelength, S is the total number of absorber types. $\varepsilon_j(\lambda_i)$ and $c_j(r)$ are the molar extinction coefficient and concentration of the absorber type j at wavelength λ_j. In our simulation experiment, we have simulated the ideal light fluence maps at different wavelengths from the absorber images, i.e., $\Phi(\lambda_i, r)$ is known. To perform spectral un-mixing, we can normalize the total light absorption by its corresponding light fluence via

$$\frac{h(\lambda_i, r)}{\Phi(\lambda_i, r)} = \sum_{j=1}^{S}(\varepsilon_j(\lambda_i)c_j(r)). \tag{10}$$

Letting $H = \frac{h(\lambda_i, r)}{\Phi(\lambda_i, r)}$ be the light absorption matrix of each location dependent on wavelength, $E = [\varepsilon_1 \quad \varepsilon_2...\varepsilon_s]^T$ be the vector containing the molar extinction coefficient of all the absorbers, and $C = [c_1 \quad c_2...c_s]^T$ be the concentration of each absorber, the above un-mixing problem can be expressed as

$$H = CE. \tag{11}$$

Then, the relative absorber concentration can be solved by minimizing the least square problem $\widehat{C} = \underset{C}{\text{argmin}}\|CE - \widehat{H}\|_2^2$ corresponding to Eq. (11).

In the animal imaging experiment, without an accurate measurement of the distribution of absorption and scattering coefficients throughout the imaging region, it is difficult to obtain the correct light fluence distribution in living biological tissues. Therefore, we present the spectral un-mixing results of the animal experiments solved by linear un-mixing in the supplementary materials.

4.7. Evaluation metrics

To quantitatively evaluate the proposed algorithm, we use four popular metrics in this study, including the root mean square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity index measure (SSIM) to evaluate the quality of the reconstructed images and Dice coefficient (DICE) to indicate the accuracy of spectral un-mixing:

$$RMSE = \sqrt{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(R_{ij} - x_{ij})^2}, \tag{12}$$

$$PSNR = 10 \cdot \log_{10}\left(\frac{L^2}{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(R_{ij} - x_{ij})^2}\right), \tag{13}$$

$$SSIM = \frac{(2\mu_R\mu_x + C_1)(2\sigma_{R \cdot x} + C_2)}{(\mu_R^2 + \mu_x^2 + C_1)(\sigma_R^2 + \sigma_x^2 + C_2)},$$ (14)

$$DICE = \frac{2\sum_{i=1}^{m}\sum_{j=1}^{n}(R_{ij} \cap x_{ij})}{\sum_{i=1}^{m}\sum_{j=1}^{n}(R_{ij} + x_{ij})},$$ (15)

where R and x are the reference image and the corresponding reconstructed image, which have the size $m \times n$. $(\cdot)_{ij}$ denotes the $i-th$ row and $j-th$ column element. L is the dynamic range of the pixel value ($L = 255$ for 8 bits grayscale images). μ_I and μ_x are the mean value of I and x. σ_I^2 and σ_x^2 are the corresponding variance, respectively. $\sigma_{I \cdot x}$ is the covariance of I and x. The small constants C_1 and C_2 are given by, $C_1 = (K_1L)^2$, $C_2 = (K_2L)^2$, where K_1, K_2 are set as $K_1 = 0.01$ and $K_2 = 0.03$, respectively.

5. Results

5.1. Simulation results

In the simulation experiments, we use the simulated interlaced sparse sampled photoacoustic signal as the input for image reconstruction. We compare the differences between the proposed method and other model-based and DL-based methods at different sparse sampling rates.

Fig. 5 shows the reconstruction results at 800 nm. MB-TV and MB-dTV are the model-based methods with TV and dTV regularization, respectively. Moreover, an end-to-end U-Net model is trained using the same data for comparison to our NDU model. As can be seen in Fig. 5, at 1/4 sparse sampling rate, all the methods exhibit fine performance, except for MB-TV whose result appears blurry due to excessive

smoothing. When shifting to 1/6 sparse sampling rate, the model-based methods are impacted by artifacts. Even though U-Net is able to generate good structural information, obvious artifacts are present. NDU-dTV produces superior performance compared to NDU-TV, owing to the incorporation of structural prior constraints.

Table 1 presents the quantitative comparison results for different reconstruction methods. As can be seen, the NDU-dTV method achieves superior metrics at the two sparse sampling rates. At 1/4 sampling, the RMSE values of the NDU-dTV method are reduced by 68.40 %, 52.26 %, 40.69 % and 40.83 % compared with MB-TV, MB-dTV, U-Net, and NDU-TV respectively. Furthermore, because of the incorporation of structural prior information as a constraint, NDU-dTV achieves the optimal SSIM value (0.9034) compared to other methods, showcasing a more complete structural reconstruction capability.

Fig. 6 presents the image intensity profile corresponding to the solid red line position in Fig. 5. As can be seen, the pixel intensity of the MB-TV (green curve) and MB-dTV (purple curve) methods are significantly lower than the reference image (black curve) at the two sparse sampling

Table 1
Quantitative comparison results of the simulation experiment at 800 nm. The optimal results are denoted by bolded values.

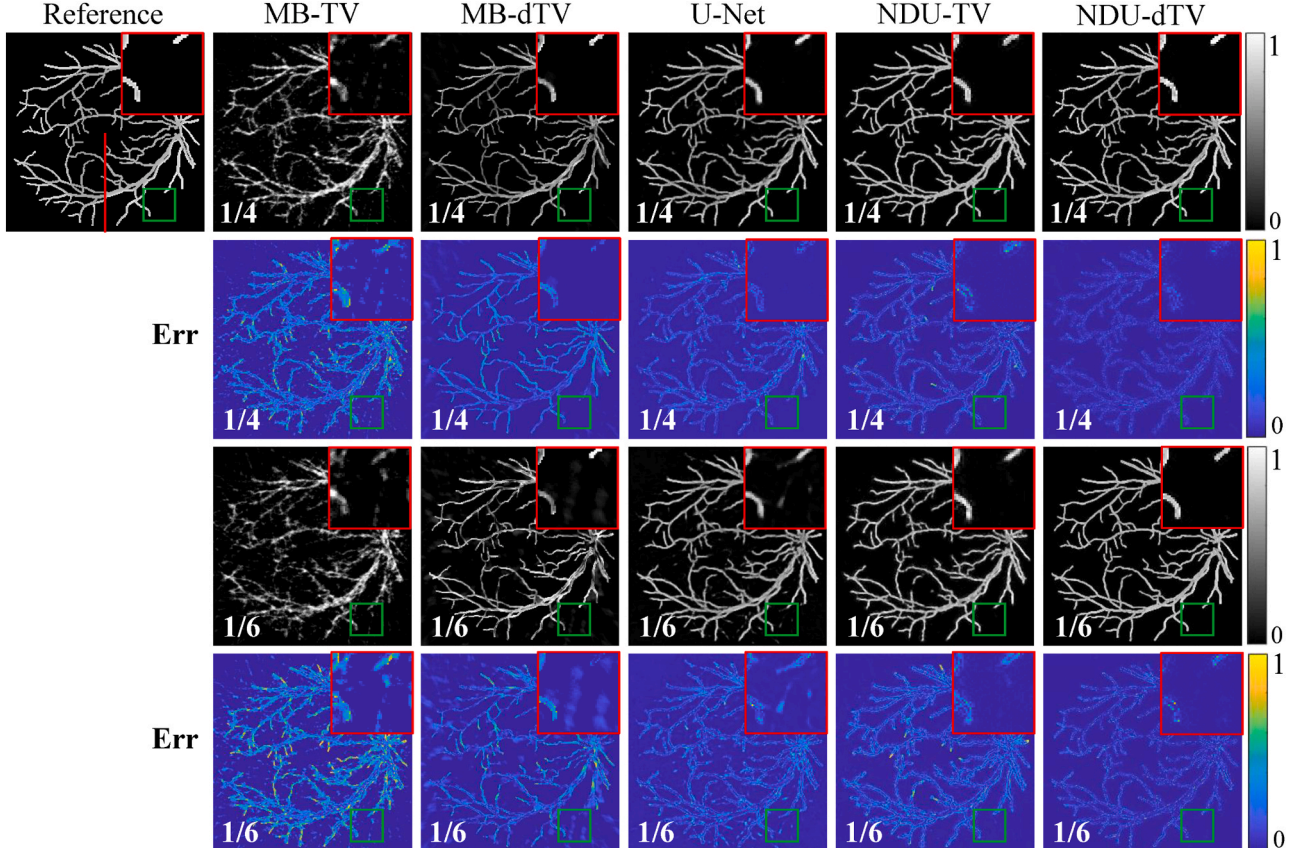| Sampling Rate	1/4			1/6		
Metrics	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
MB-TV	0.1573	18.0675	0.6218	0.1706	16.3973	0.5170
MB-dTV	0.1041	19.6528	0.8606	0.1146	18.8169	0.7522
U-Net	0.0838	22.1024	0.8577	0.0991	20.7787	0.8281
NDU-TV	0.0840	22.5148	0.8139	0.1121	20.0046	0.6377
NDU-dTV	**0.0497**	**26.0806**	**0.9034**	**0.0729**	**22.7506**	**0.8449**

Fig. 5. Visual comparison of reconstructed images by different methods. '1/4', '1/6' denote the sparse sampling rates. Err shows the difference between the reconstructed image and the reference image. The artifact reduction performance of the different methods is compared by the enlarged red boxes.
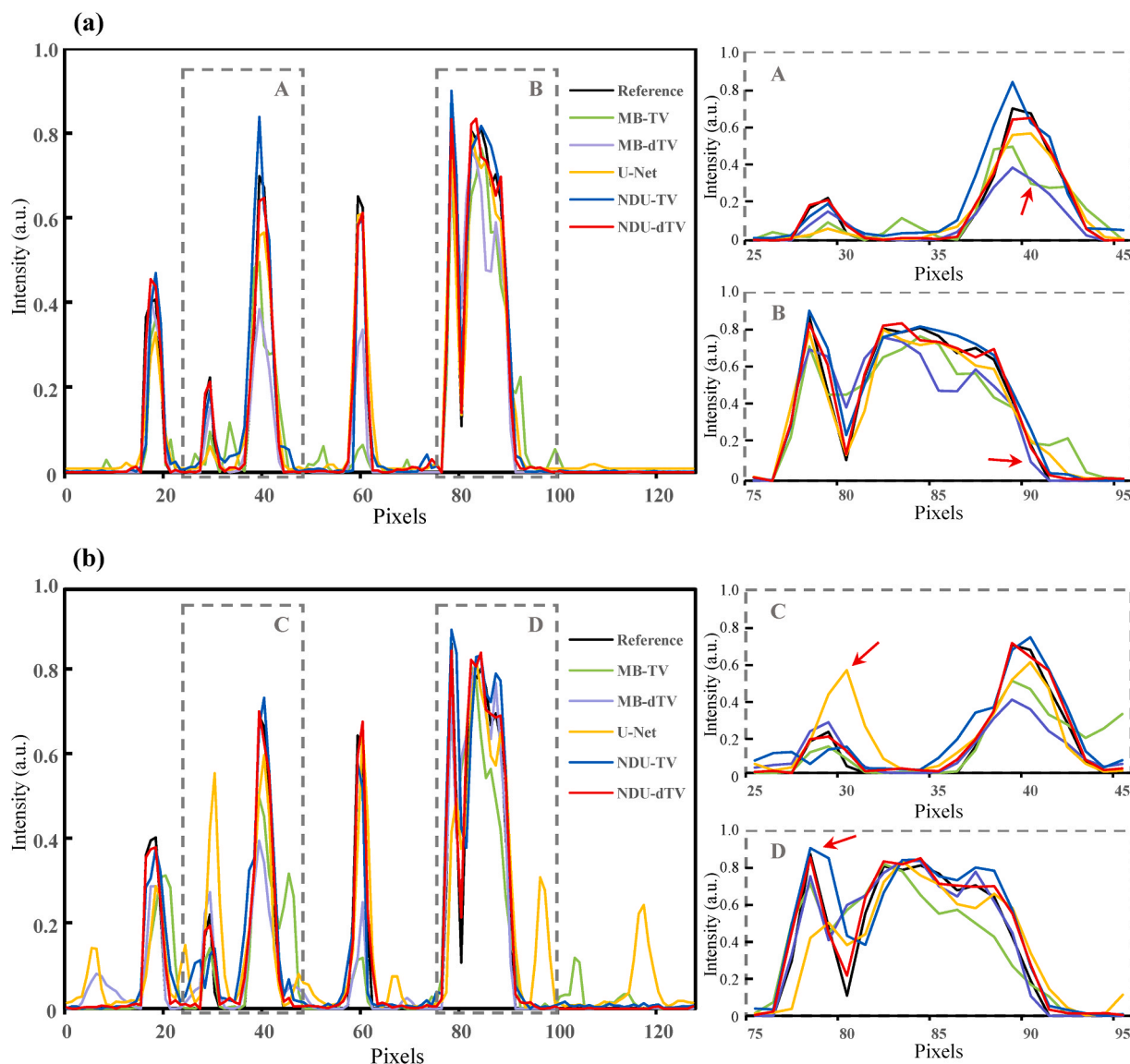
Fig. 6. Image intensity profiles corresponding to the solid red line position in Fig. 5 at 1/4 down-sampling rate (a) and at 1/6 down-sampling rate (b). The dashed-box regions are enlarged on the right panel for better comparison.

rates, with larger errors indicated by the red arrows. While U-Net demonstrates commendable performance at 1/4 sampling rate, its efficacy diminishes notably at 1/6 sampling rate. This can be found on the profile position indicated by the arrow in Fig. 6(C), where oddly intense signal is presented. Without the constraint of structural prior information, broadening of the edges is observed in NDU-TV, as in Fig. 6(D). The NDU-dTV method (red curve) has a smoother profile, and the pixel values are closer to the reference profile, as indicated by the arrow in Fig. 6(B).

Next, we conduct spectral un-mixing experiments at five different wavelengths. Fig. 7 shows the spectral un-mixing results of different methods at 1/6 sampling rate. We define the absorber in the left half of the image as HbO2 and in the right half as Hb for better comparison. It is obvious that the NDU-dTV method produces superior un-mixing results. Clear and complete absorber distribution has been revealed that resembles the reference image closely. For quantitative comparison, we subsequently calculate the DICE of the un-mixing results. As shown in Fig. 8, NDU-dTV achieves the highest DICE score. In particular, the DICE score of HbO_2 of NDU-dTV is 56.61 %, 19.41 %, 5.89 % and 8.46 % higher than those of MB-TV, MB-dTV, U-Net, and NDU-TV, respectively, whereas the DICE score of Hb is 45.50 %, 17.36 %, 6.48 % and 9.16 %

higher.

Fig. 9 presents the spectral un-mixing intensity profiles corresponding to the solid yellow line position in Fig. 7. As can be seen, the pixel intensity of the MB-TV (green curve) and MB-dTV (purple curve) methods exhibit notable deviations from the reference image (black curve), as evident in the error maps presented in Fig. 7. Introducing the deep algorithm unrolling, the NDU-TV gets a notable improvement in the image reconstruction, leading to more accurate intensity in un-mixing results compared to the MB-TV. With the constraint of structural prior information, the NDU-dTV has the optimal results in the un-mixing intensity profiles.

5.2. Small animal imaging experiment results

Fig. 10 shows the imaging results of the abdomen region of the mouse using different image reconstruction methods at 1/4 down-sampling rate at 800 nm. As can be seen, the MB-TV methods are affected by artifacts that resulted in blurry internal organs. By introducing the structural prior constraints, the image quality is improved by the MB-dTV method. Although the U-Net can effectively eliminate artifacts, it inadvertently erases organ detail information, as indicated by
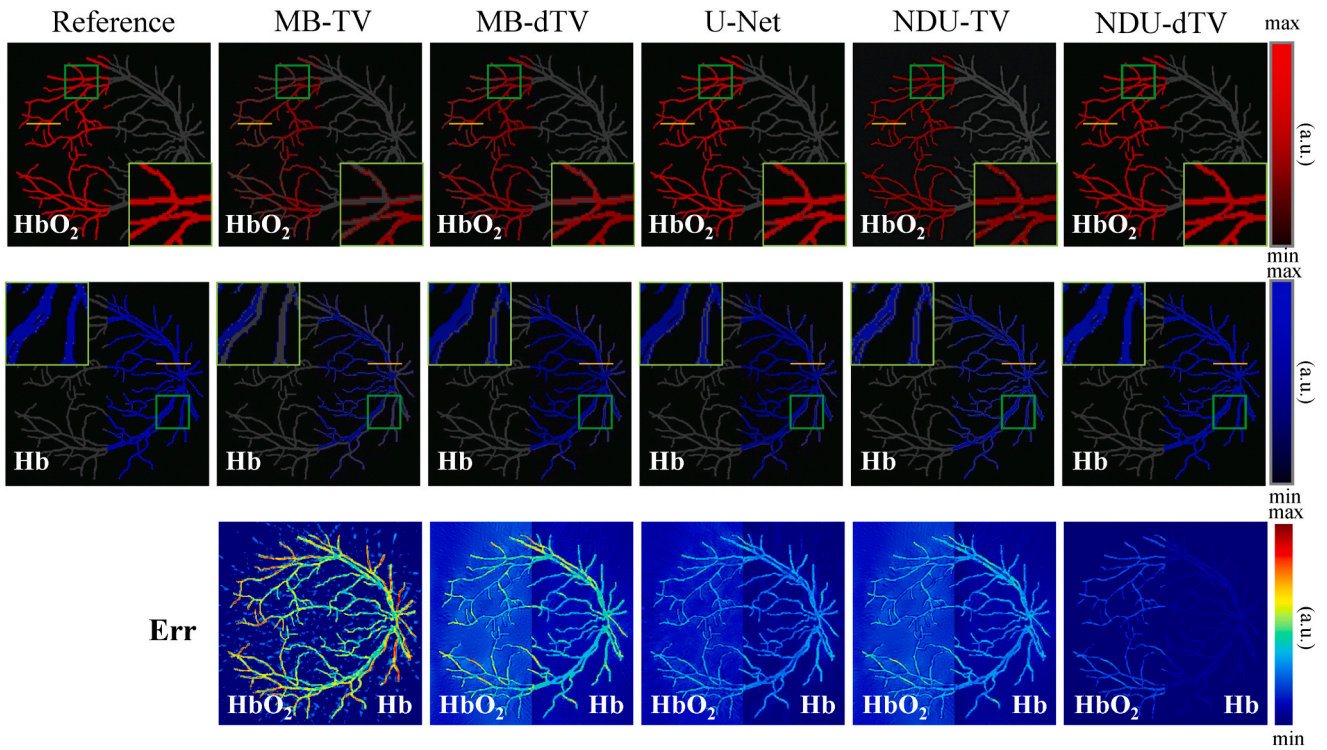
Fig. 7. Spectral un-mixing results of different methods at 1/6 sampling rate. The absorbers are HbO_2 (red) and Hb (blue). The enlarged areas corresponding to the green boxes for better visualization. Err shows the difference between the spectral un-mixing images and the reference image.
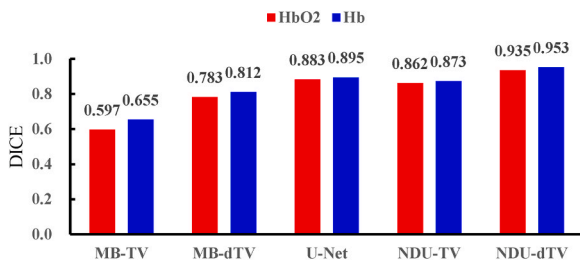

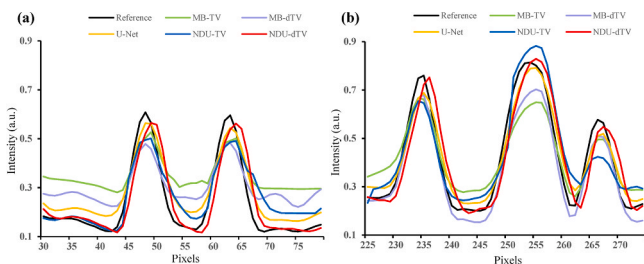**Fig. 8.** The DICE score of the spectral un-mixing results in Fig. 7.

Fig. 9. HbO_2 (a) and Hb (b) image intensity profiles corresponding to the solid yellow line position in Fig. 7 at 1/6 down-sampling rate.

the red arrow. With the advantage provided by the DAU method, the image quality has been improved significantly by NDU-TV. Compared to NDU-TV which made a rough balance between image denoising performance and detail structure preservation, the proposed NDU-dTV method achieves more comprehensive reconstruction of detailed image structures.

Table 2 lists quantitative evaluation results in the small animal imaging experiments. The NDU-dTV method has the best results at both 1/4 and 1/6 down-sampling rates. In addition, due to the introduction of

structural prior information as a constraint, the NDU-dTV method is able to present a more complete structure with less error, and its RMSE value is the lowest among all methods. Specifically, at the 1/6 down-sampling rate, the RMSE values of the NDU-dTV method are reduced by 56.83 %, 41.86 %, 19.01 %, and 38.04 % when compared to MB-TV, MB-dTV, U-Net, and NDU-TV, indicating that the NDU-dTV method has better reconstruction results.

In addition, Fig. 11 shows the intensity profiles corresponding to the solid line position in Fig. 10 for different methods. As can be seen from the profiles at 1/6 down-sampling rate, the signal curves of MB-TV (green curve) and MB-dTV (purple curve) exhibits notable deviations from the reference curve (black curve) indicated by Fig. 11(C). Compared to the NDU-TV (blue curve), the NDU-dTV (red curve) yields a smoother profile with pixel intensity more closely approximating the reference profile, as depicted in Fig. 11(A).

5.3. Model structure selection

The optimized hyperparameters are closely associated with the DAU architecture, including the number of stages (K) in the image restoration block and the number of iterations (N) in ADMM. We calculate the NMSE between the output of the image reconstruction layer and the reference image at 1/6 sampling rate. As shown in Fig. 12(a), the NMSE decreases fast when the number of stages $K \leq 10$. Fig. 12(b) shows the evaluation of the number of ADMM iterations, i.e. N, given $K = 10$. As can be seen, N can be set to 10 considering the reconstruction quality and computational efficiency.

5.4. Computational efficiency results

To further assess the computational performance, we also calculate the parameters and processing time of different methods. As shown in Table 3, while maintaining superior image quality, the reconstruction time required by the proposed method is, on average, 19.0 times less than that of the model-based approach. Additionally, NDU-dTV also
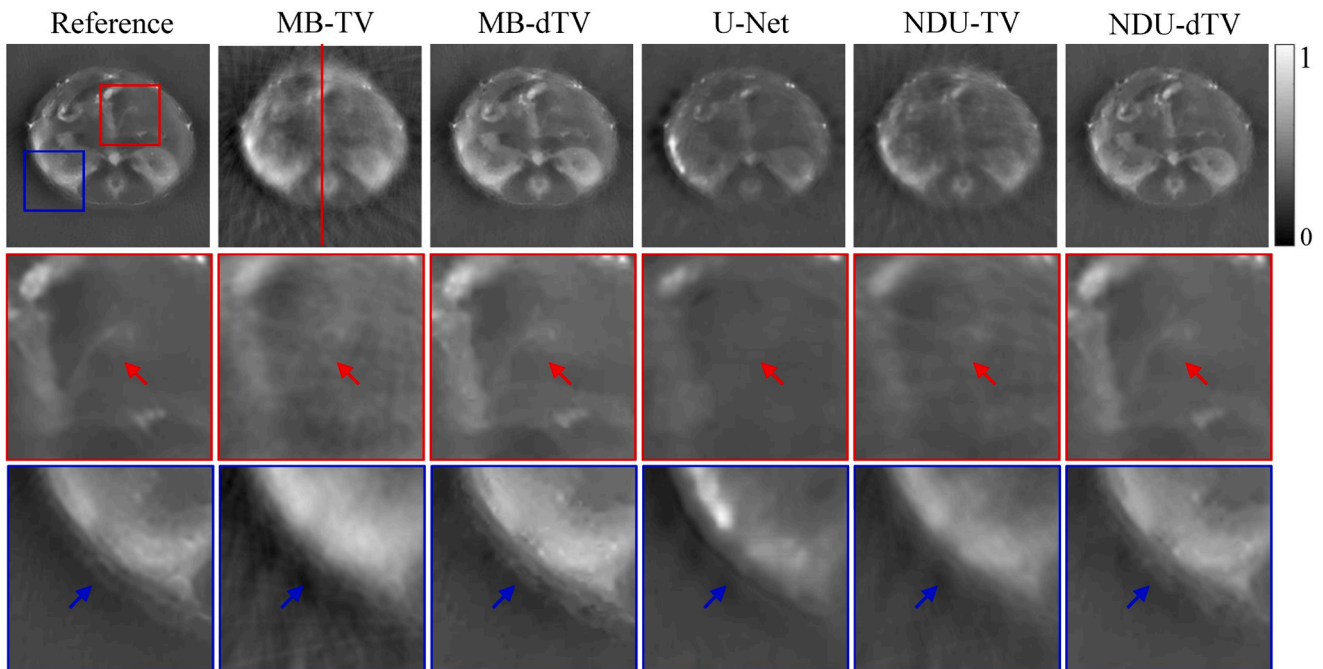
Fig. 10. *In vivo* mouse image reconstruction results by different methods at 1/4 sparse sampling rate at 800 nm. The second and third rows are enlarged areas corresponding to the red and blue boxes in the first row for better visualization.

Table 2
Quantitative evaluation of the small animal imaging results. The optimal results are denoted by bolded values.

Sampling Rate	1/4			1/6		
Metrics	RMSE	PSNR	SSIM	RMSE	PSNR	SSIM
MB-TV	0.0945	20.4958	0.8299	0.1332	17.5753	0.7393
MB-dTV	0.0567	24.9244	0.9270	0.0989	20.0923	0.8697
U-Net	0.0617	24.1950	0.8641	0.0710	22.9765	0.8492
NDU-TV	0.0763	22.3539	0.8369	0.0928	20.6521	0.8190
NDU-dTV	**0.0432**	**27.2890**	**0.9512**	**0.0575**	**24.8018**	**0.9154**

prevails the U-Net method.

5.5. Effectiveness of different image restoration blocks

Finally, we conduct quantitative experiments to assess the contribution of the image restoration blocks. Table 4 shows the RMSE indices at 1/6 sparse sampling rate for the simulation experiments after implementing different image restoration blocks, where "TV" and "dTV" indicate the image restoration block based on off-the-shelf TV and dTV denoiser, representing the standard DAU model. "DAU-TV" and "DAU-dTV" denote the image restoration block based on DAU, representing the nested DAU models. As can be seen, owing to the adaptive learning of suitable prior, the nested DAU method outperforms the standard DAU method, further proving the effectiveness of the nested DAU framework.

6. Discussion

The above rigorous experimental results show that the proposed DAU-dTV method produce image reconstruction results with overall quality and spectral un-mixing accuracy comparable to dense sampling. Compared to model-based method that requires handcrafted regularization, our DAU-based method automatically learns the suitable prior from the training dataset. Compared to previous DL-based PAT image reconstruction methods, our network architecture takes into account explicit domain knowledge, including both the physical imaging model

and the sparse sampling setting. Moreover, the DAU framework, as the core of our method, has been shown to possess the advantage of good interpretability over conventional black-box-like DL models [23,47]. Our method shares the same advantage of DAU.

Although the DAU method has made success in PAT image reconstruction, it is unable to introduce structural prior information of sparse sampling PAT. To alleviate the deficiency, our NDU-dTV incorporates the prior information through PnP-ADMM and achieves enhanced performance and flexibility. Instead of using pre-train denoiser, we integrate the dTV algorithm, which considers local directional image features, into DAU to serve as an image restoration block. The experimental results have demonstrated improved image reconstruction performance closely associates with the unique structure of the proposed NDU-dTV model. Compared to standard DAU, our NDU-dTV show lower RMSE, indicating high feasibility. Compared to conventional U-Net model, we achieve significantly enhanced image quality with three times fewer learnable parameters. While maintaining optimal performance, the reconstruction speed of our method is also much faster.

We evaluate the performance of our NDU-dTV method using the interlaced sparse sampling scheme proposed previously. This scheme has certain difference compared to conventional sparse sampling method because a clean structural prior image is obtained during multispectral imaging. However, this does not affect the applicability of our method to other sparse sampling strategy, such as non-interlaced sparse sampling or limited-view sampling. It can be confirmed by the NDU-TV results in both the simulation and real-world experiments, where no structural prior information is introduced.

7. Conclusion

In this work, we propose NDU-dTV, a novel deep-learning-based PAT image reconstruction method based on deep algorithm unrolling. Developed upon classical ADMM optimization framework, our method integrates the dTV algorithm into the DAU process for structural information introduction. Our NDU-dTV incorporates both the adaptive parameter optimization of DL-based method and the interpretability of model-based method. We conduct extensive experiments of sparse sampling PAT imaging to evaluate the performance of NDU-dTV. The
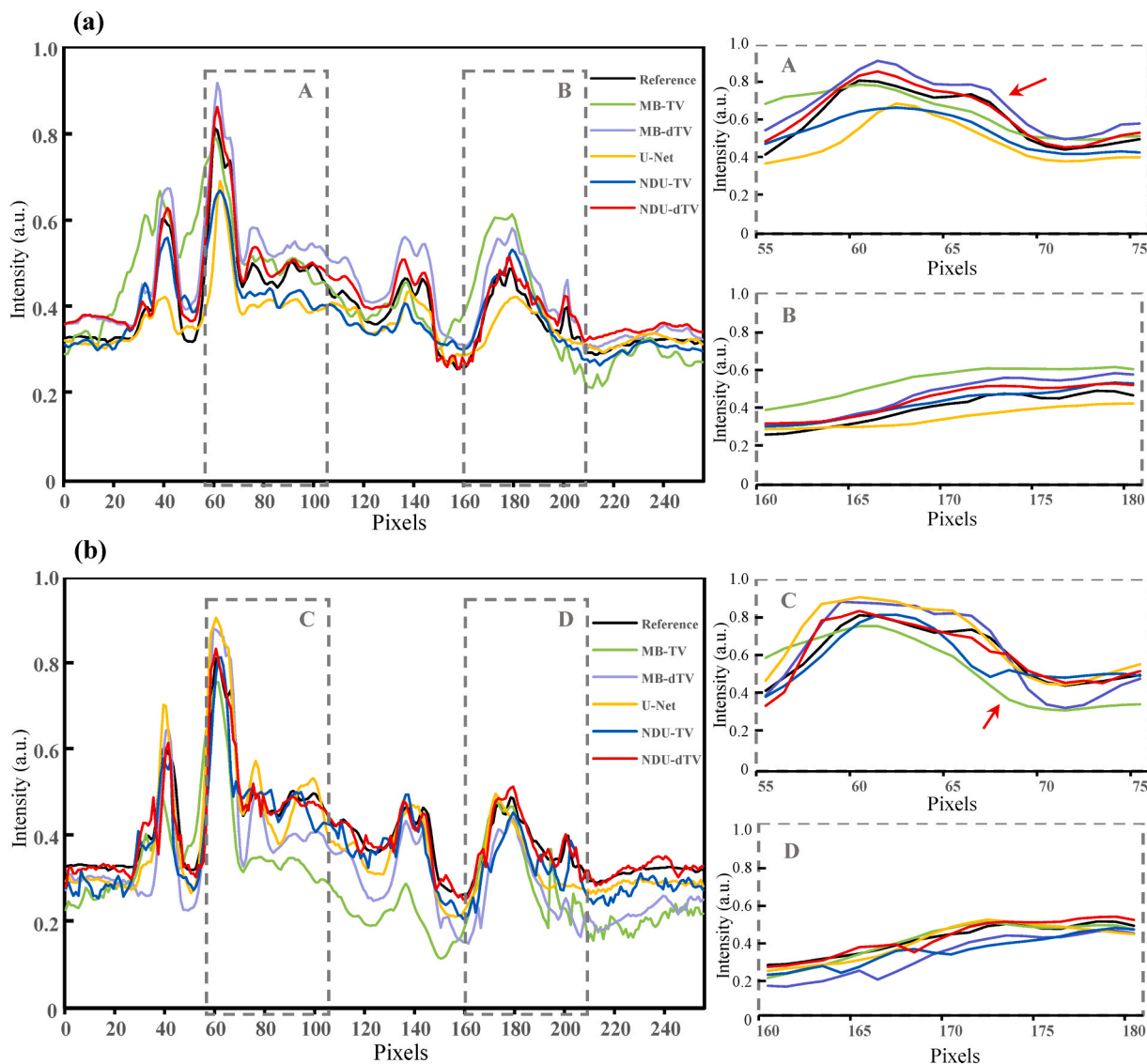
Fig. 11. Image intensity profiles corresponding to solid red line position in Fig. 10 at 1/4 down-sampling rate (a) and 1/6 down-sampling rate (b). The dashed-box regions are enlarged on the right panel for better comparison.
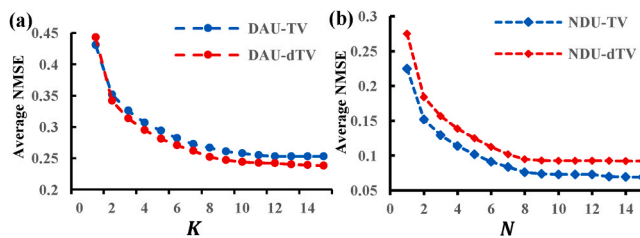

Fig. 12. (a) Average NMSEs using different number of stages (*K*) in image restoration block. (b) Average NMSEs using different number of iterations (*N*) in ADMM given *K*=10.

results demonstrate our method offers optimal image reconstruction quality with high computational efficiency and fine interpretability.

CRediT authorship contribution statement

Jia Ge: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. **Zongxin Mo**: Writing – original draft, Visualization, Validation, Software, Methodology,

Table 3

Computational performance comparison on *in vivo* animal experiment.

Method	MB-TV	MB-dTV	U-Net	NDU-TV	NDU-dTV
Parameters	N/A	N/A	3.22×10^5	1.02×10^5	1.02×10^5
Testing Time	116.577 s	127.322 s	1.151 s	5.852 s	6.651 s
Training Time	N/A	N/A	9 h	16 h	17 h

Table 4

Comparison of different image restoration block settings.

Image restoration block	TV	dTV	DAU-TV	DAU-dTV
RMSE	0.1428	0.0921	0.1083	0.0746

Investigation, Formal analysis, Data curation. **Shuangyang Zhang**: Investigation, Data curation. **Xiaoming Zhang**: Investigation. **Yutian Zhong**: Data curation. **Zhaoyong Liang**: Visualization. **Chaobin Hu**: Investigation. **Wufan Chen**: Supervision, Funding acquisition. **Li Qi**: Writing – original draft, Supervision, Resources, Project administration,

Funding acquisition, Conceptualization.

Declaration of Competing Interest

Data availability

Data will be made available on request.

Acknowledgements

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.pacs.2024.100618.

References

[1] D. Razansky, M. Distel, C. Vinegoni, R. Ma, N. Perrimon, R.W. Köster, V. Ntziachristos, Multispectral opto-acoustic tomography of deep-seated fluorescent proteins in vivo, Nat. Photonics 3 (2009) 412–417.

[2] X. Li, S. Zhang, J. Wu, S. Huang, Q. Feng, L. Qi, W. Chen, Multispectral interlaced sparse sampling photoacoustic tomography, IEEE Trans. Med. Imaging 39 (2020) 3463–3474.

[3] S. Gottschalk, O. Degtyaruk, B. Mc Larney, J. Rebling, M.A. Hutter, X.L. Dean-Ben, S. Shoham, D. Razansky, Rapid volumetric optoacoustic imaging of neural dynamics across the mouse brain, Nat. Biomed. Eng. 3 (2019) 392–401.

[4] S. Zhang, Z. Liang, K. Tang, X. Li, X. Zhang, Z. Mo, J. Wu, S. Huang, J. Liu, Z. Zhuang, L. Qi, W. Chen, In vivo co-registered hybrid-contrast imaging by successive photoacoustic tomography and magnetic resonance imaging, Photoacoustics 31 (2023) 100506.

[5] Z. Liang, S. Zhang, J. Wu, X. Li, Z. Zhuang, Q. Feng, W. Chen, L. Qi, Automatic 3-D segmentation and volumetric light fluence correction for photoacoustic tomography based on optimal 3-D graph search, Med. Image Anal. 75 (2022) 102275.

[6] K. Basak, X. Luis Dean-Ben, S. Gottschalk, M. Reiss, D. Razansky, Non-invasive determination of murine placental and foetal functional parameters with multispectral optoacoustic tomography, Light, Sci. Appl. 8 (2019) 71.

[7] S. Zhang, J. Liu, Z. Liang, J. Ge, Y. Feng, W. Chen, L. Qi, Pixel-wise reconstruction of tissue absorption coefficients in photoacoustic tomography using a non-segmentation iterative method, Photoacoustics 28 (2022) 100390.

[8] M. Xu, L.V. Wang, Universal back-projection algorithm for photoacoustic computed tomography, Phys. Rev. E Stat. Nonlin Soft Matter Phys. 71 (2005) 016706.

[9] Y. Wang, T. Lu, J. Li, W. Wan, W. Ma, L. Zhang, Z. Zhou, J. Jiang, H. Zhao, F. Gao, Enhancing sparse-view photoacoustic tomography with combined virtually parallel projecting and spatially adaptive filtering, Biomed. Opt. Express 9 (2018) 4569–4587.

[10] L. Yao, H. Jiang, Photoacoustic image reconstruction from few-detector and limited-angle data, Biomed. Opt. Express 2 (2011) 2649–2654.

[11] K. Tang, S. Zhang, Z. Liang, Y. Wang, J. Ge, W. Chen, L. Qi, Advanced Image Post-Processing Methods for Photoacoustic Tomography: A Review. in: Photonics, MDPI, 2023, p. 707.

[12] A. Buehler, A. Rosenthal, T. Jetzfellner, A. Dima, D. Razansky, V.J. Mp Ntziachristos, Model-based optoacoustic inversions with incomplete projection data, Med. Phys. 38 (2011) 1694–1704.

[13] X. Li, J. Ge, S. Zhang, J. Wu, L. Qi, W. Chen, Multispectral interlaced sparse sampling photoacoustic tomography based on directional total variation, Comput. Methods Prog. Biomed. 214 (2022) 106562.

[14] A. Taruttis, V. Ntziachristos, Advances in real-time multispectral optoacoustic imaging and its applications, Nat. Photonics 9 (2015) 219–227.

[15] J. Wang, Y. Wang, Photoacoustic imaging reconstruction using combined nonlocal patch and total-variation regularization for straight-line scanning, Biomed. Eng. OnLine 17 (2018) 1–24.

[16] Y. Zhang, Y. Wang, C. Zhang, Total variation based gradient descent algorithm for sparse-view photoacoustic image reconstruction, Ultrasonics 52 (2012) 1046–1055.

[17] X. Li, L. Qi, S. Zhang, S. Huang, J. Wu, L. Lu, Y. Feng, Q. Feng, W. Chen, Model-based optoacoustic tomography image reconstruction with non-local and sparsity regularizations, IEEE Access 7 (2019) 102136–102148.

[18] K. Tang, S. Zhang, Y. Wang, X. Zhang, Z. Liu, Z. Liang, H. Wang, L. Chen, W. Chen, L. Qi, Learning spatially variant degradation for unsupervised blind photoacoustic tomography image restoration, Photoacoustics 32 (2023) 100536.

[19] Z. Liang, S. Zhang, A. Wei, Z. Liu, Y. Wang, H. Hu, W. Chen, L. Qi, CylinGCN: cylindrical structures segmentation in 3D biomedical optical imaging by a contour-based graph convolutional network, Comput. Med. Imaging Graph. (2023) 102316.

[20] W. Li, Z. Huang, C. Zhou, X. Zhang, W. Fan, D. Liang, Z. Hu, Adaptive 3D noise level-guided restoration network for low-dose positron emission tomography imaging, Interdiscip. Med. (2023) e20230012.

[21] D. Waibel, J. Gröhl, F. Isensee, T. Kirchner, K. Maier-Hein, L. Maier-Hein, Reconstruction of Initial Pressure from Limited View Photoacoustic Images Using Deep Learning. in: Photons Plus Ultrasound: Imaging and Sensing 2018, SPIE, 2018, pp. 196–203.

[22] H. Lan, D. Jiang, C. Yang, F. Gao, F. Gao, Y-Net: Hybrid deep learning image reconstruction for photoacoustic tomography in vivo, Photoacoustics 20 (2020) 100197.

[23] V. Monga, Y. Li, Y.C. Eldar, Algorithm unrolling: interpretable, efficient deep learning for signal and image processing, IEEE Signal Process. Mag. 38 (2021) 18–44.

[24] C. Mou, Q. Wang, J. Zhang, Deep Generalized Unfolding Networks for Image Restoration, 2022 IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (2022) 17378–17389.

[25] K. Gregor, Y. LeCun, Learning fast approximations of sparse coding, Proc. 27th Int. Conf. Int. Conf. Mach. Learn. (2010) 399–406.

[26] Y. Li, M. Tofighi, J. Geng, V. Monga, Y.C. Eldar, Efficient and interpretable deep blind image deblurring via algorithm unrolling, IEEE Trans. Comput. Imaging 6 (2020) 666–681.

[27] M. Nagahama, K. Yamada, Y. Tanaka, S.H. Chan, Y.C. Eldar, Graph signal restoration using nested deep algorithm unrolling, IEEE Trans. Signal Process. 70 (2022) 3296–3311.

[28] J. Schwab, S. Antholzer, M. Haltmeier, Learned Backprojection for Sparse and Limited View Photoacoustic Tomography. in: Photons Plus Ultrasound: Imaging and Sensing 2019, SPIE, 2019, pp. 263–271.

[29] 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 Trans. Med. Imaging 37 (2018) 1382–1393.

[30] Y. Yang, J. Sun, H. Li, Z. Xu, Deep ADMM-Net for compressive sensing MRI, Adv. Neural Inf. Process. Syst. (2016) 29.

[31] A. Dima, N.C. Burton, V.J.Jo.B.O. Ntziachristos, Multispectral optoacoustic tomography at 64, 128, and 256 channels, J. Biomed. Opt. 19 (2014).

[32] A. Beck, M.J.Sjois Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (2009) 183–202.

[33] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends® Mach. Learn. 3 (2011) 1–122.

[34] A. Chambolle, T.J.Jomi Pock, vision, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imaging 40 (2011) 120–145.

[35] M. Hong, Z.-Q. Luo, M. Razaviyayn, Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems, SIAM J. Optim. 26 (2016) 337–364.

[36] K. Wei, A. Aviles-Rivero, J. Liang, Y. Fu, C.-B. Schönlieb, H. Huang, Tuning-free plug-and-play proximal algorithm for inverse imaging problems, Int. Conf. Mach. Learn. PMLR (2020) 10158–10169.

[37] J. He, Y. Yang, Y. Wang, D. Zeng, Z. Bian, H. Zhang, J. Sun, Z. Xu, J. Ma, Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction, IEEE Trans. Med. Imaging 38 (2019) 371–382.

[38] S.H. Chan, Performance analysis of plug-and-play ADMM: a graph signal processing perspective, IEEE Trans. Comput. Imaging 5 (2019) 274–286.

[39] R. Hou, F. Li, G. Zhang, Truncated residual based plug-and-play ADMM algorithm for MRI reconstruction, IEEE Trans. Comput. Imaging 8 (2022) 96–108.

[40] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, R.J.I.To.P.A. Timofte, M, Intelligence, Plug-and-play image restoration with deep denoiser prior, IEEE Trans. Pattern Anal. Mach. Intell. 44 (2021) 6360–6376.

[41] J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, B. van Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE Trans. Med. Imaging 23 (2004) 501–509.

[42] A. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, IEEE Trans. Med. Imaging 19 (2000) 203–210.

[43] N. Bosschaart, G.J. Edelman, M.C. Aalders, T.G. van Leeuwen, D.J. Faber, A literature review and novel theoretical approach on the optical properties of whole blood, Lasers Med Sci. 29 (2014) 453–479.

[44] M. Schweiger, S. Arridge, The Toast++ software suite for forward and inverse modeling in optical tomography, J. Biomed. Opt. 19 (2014).

[45] L. Qi, S. Huang, X. Li, S. Zhang, L. Lu, Q. Feng, W. Chen, Cross-sectional photoacoustic tomography image reconstruction with a multi-curve integration model, Comput. Methods Prog. Biomed. 197 (2020) 105731.

[46] F. Ozdemir, B. Lafci, X.L. Deán-Ben, D. Razansky, F. Perez-Cruz, OADAT: experimental and synthetic clinical optoacoustic data for standardized image processing, IEEE Trans. Mach. Learn. Res. (2023).

[47] Y. Yang, J. Sun, H. Li, Z. Xu, ADMM-CSNet: a deep learning approach for image compressive sensing, IEEE Trans. Pattern Anal. Mach. Intell. 42 (2020) 521–538.

Jia Ge received his master's degree from Southern Medical University. His research interests include photoacoustic imaging and image reconstruction.

Zhaoyong Liang received his bachelor's degree from Southern Medical University in 2022. He is currently pursuing the master's degree in the school of Biomedical Engineering, Southern Medical University. His research interests include photoacoustic tomography and medical image processing.

Zongxin Mo received his bachelor's degree from Guangzhou Medical University in 2022. He is currently pursuing the master's degree at the School of Biomedical Engineering, Southern Medical University. His research interests include photoacoustic tomography and medical image processing.
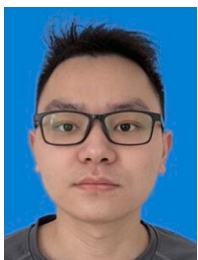
Chaobin Hu received his bachelor's degree from SouthernMedical University in 2023. He is currently pursuing the master's degree in the school of Biomedical Engineering, Southern Medical University. His research interests include deep learning-based photoacoustic tomography image reconstruction.

Shuangyang Zhang received his Ph.D. degree in 2022 from Southern Medical University. He is currently a postdoctoral fellow at the Guangdong Provincial Key Laboratory of Medial Image Processing, School of Biomedical Engineering, Southern Medical University in Guangzhou, China. His research interests include photoacoustic imaging and multimodal image registration.

Wufan Chen received his B.S. degree in 1975 and M.Sc. degree in 1981, both from Beihang University. He is currently a full professor at the Guangdong Provincial Key Laboratory of Medial Image Processing, School of Biomedical Engineering, Southern Medical University. His research interests include biomedical imaging principle and image processing.

Xiaoming Zhang received his bachelor's degree from Southern Medical University in 2022. He is currently pursuing the master's degree in the school of Biomedical Engineering, Southern Medical University. His research interests include image restoration based on photoacoustic Tomography and relevant system based on photoacoustic imaging and MRI.

Li Qi received his Ph.D. degree in Optical Engineering in 2016 from Nanjing University. He is currently an associate professor at the Guangdong Provincial Key Laboratory of Medial Image Processing, School of Biomedical Engineering, Southern Medical University in Guangzhou, China. His research interests include photoacoustic imaging and optical coherence tomography.

Yutian Zhong received his bachelor's degree from Southern Medical University in 2022. He is currently pursuing the master's degree in the school of Biomedical Engineering, Southern Medical University. His research interests include photoacoustic tomography and medical image processing.