Self-supervised light fluence correction network for photoacoustic tomography based on diffusion equation

Zhaoyong Liang [a,b,c], Zongxin Mo [a,b,c], Shuangyang Zhang [a,b,c], Long Chen [a,b,c], Danni Wang [a,b,c], Chaobin Hu [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

Deep learning (DL) shows promise in estimating the absorption coefficient distribution of biological tissue in quantitative photoacoustic tomography (QPAT) imaging, but its application is limited by a lack of ground truth for supervised network training. To address this issue, we propose a DL-based light fluence correction method that only uses the original PAT images for network training. Our self-supervised QPAT network model, which we termed SQPA-Net, introduces light fluence estimation based on diffusion equation to the loss function, and thus guides the model to learn an implicit representation of photoacoustic light transport within tissue. Simulation and small animal imaging experiments demonstrate the effectiveness and efficiency of our method. Compared to current DL-based methods and traditional iterative correction method, the proposed SQPA-Net achieves better light fluence correction results and significantly reduces the processing time.

1. Introduction

Photoacoustic tomography (PAT) combines the advantages of optical and acoustic imaging, achieving optical contrast images while also providing greater imaging depth [1–4]. Multispectral PAT can perform multispectral imaging using laser light of different wavelengths [5–8], enabling the differentiation of the distribution of endogenous tissue absorbers, such as oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb), and is applied in various preclinical studies [9,10] and clinical trials [11,12].The initial photoacoustic maps do not accurately reflect the optical properties of various tissues, as PAT can be viewed as the result of multiplying an inhomogeneous light fluence (LF) by an absorption coefficient (μ_a) [13]. By eliminating the effects of LF, it is possible to recover an absorption coefficient map that represents the characteristics of the tissue from the original photoacoustic image, enabling quantitative analysis [14]. Therefore, it is essential to address the nonlinear optical inversion problem in quantitative PAT, or QPAT. Some researchers are attempting to tackle the optical inverse problem using other imaging techniques, such as diffuse optical tomography [15] and

acousto-optic theory [16], which directly address energy distribution. However, this approach requires additional equipment, which can be costly.

Additionally, model-based iterative correction methods are also applicable to address this issue. As shown in Fig. 1(a), by using a light transport model to generate the corresponding LF distribution and iteratively minimizing the loss function between the measured data and the estimated data, an absorption coefficient image can be obtained [17]. For instance, Cox et al. [14] were the first to combine the diffusion equation with iterative methods to solve for the absorption coefficient. Liu et al. [18] further improved the solution accuracy using Monte Carlo simulations. Zhang et al. [19] proposed a two-step iterative strategy that achieves high-precision absorption coefficient reconstruction without the need for pre-segmentation. While these methods can attain a high level of correction accuracy, they are time-intensive and fail to facilitate real-time LF correction.

Currently, deep learning methods show great potential in the field of quantitative PAT. These methods utilize extensive data for training, learning hidden mapping relationships to enable fast resolution, as

---

* Corresponding author at: School of Biomedical Engineering, Southern Medical University, 1023 Shatai Rd., Baiyun District, Guangzhou, Guangdong 510515, China.
E-mail address: qili@smu.edu.cn (L. Qi).

shown in Fig. 1(c). Among numerous neural networks, U-Net, based on convolutional structures, is the most widely used. For instance, Cai et al. [20] used residual units to estimate chromophore concentrations or blood oxygen levels, reducing result errors and enhancing image noise suppression capabilities. Luke et al. [21] employed two parallel U-Net to form an O-Net architecture, achieving sO$_2$ estimation with blood vessels segmentation. Li et al. [22] used two U-Nets to separately estimate the absorption coefficient and light fluence maps, yielding satisfactory results. In addition, other networks also play a significant role in photoacoustic quantification. Liang et al. [23] employed Fourier neural operators to learn the forward process of the light transport model and combined it with iterative methods, significantly enhancing the solving speed of the iterative correction method. Nölke et al. [24] explored the ambiguity of quantitative photoacoustic using conditional reversible neural networks, achieving complete posterior distribution outputs for sO$_2$ estimation. Wang et al. [25] proposed an extraction-attention-predictor network architecture that demonstrates higher efficiency and better quantitative results in estimating absorption coefficients. Although the above methods demonstrate good performance, they are all based on supervised training, which requires labeled datasets for implementation. However, in real-world PAT imaging, it is difficult to obtain accurate labeled light fluence map, and the high preparation costs hinder the development of deep learning in QPAT imaging. Moreover, most of the above DL-based QPAT methods are end-to-end models that do not incorporate the physical process of light-tissue interaction.

Recently, development of deep learning methods integrated with physical knowledge has been proposed in various fields. Raissi et al. [26] proposed a self-supervised learning method for solving partial differential equations by combining physical equations with loss functions. Li et al. [27] introduced Fourier neural operator, which achieves efficient representation by parameterizing integral kernels in Fourier space. Wen et al. [28] proposed a new neural network, U-FNO, which can learn multiphase flow problems in porous media, with experimental results indicating it outperforms current state-of-the-art CNN networks. Huang et al. [29] reconstructed holograms using physically consistent loss functions, significantly enhancing their generalization performance. These newly developed network models introduce regularizations based on physical principle, and thus achieve excellent results compared to traditional models.

Considering that the integration of physical information with neural networks may address the issues of traditional iterative methods and labeled-based deep learning approaches, herein we propose a self-supervised LF correction method based on the diffusion equation for quantitative PAT. Our method, which we termed SQPA-Net, achieves self-supervised network training by incorporating the optical processes in PAT imaging into the loss function through diffusion equation-based light transport model [Fig. 1(b)]. Our SQPA-Net only requires the original PAT images for network training, and once the model has been well-trained, LF correction can be performed on any input of the original PAT image and the solution speed is much faster than traditional iterative correction methods. We test the performance of SQPA-Net in simulation and small animal imaging experiment. Compared to other state-of-the-art DL-based methods, our self-supervised learning methods achieves improved LF correction results, thus demonstrating its potential in future QPAT applications.

2. Methods

2.1. The optical processes in photoacoustic imaging

The pixel values of the original PAT represent the initial pressure generated after the imaging tissue absorbs energy, which can be expressed by the following formula:

$$p(\boldsymbol{r}) = \Gamma\mu_a(\boldsymbol{r})\phi(\mu_a(\boldsymbol{r}), \mu_s(\boldsymbol{r}), g(\boldsymbol{r})), \qquad (1)$$

where $p(\boldsymbol{r})$ represents the initial pressure at point $\boldsymbol{r}$. Γ is the Gruneisen coefficient, which indicates the efficiency of conversion from thermal energy to pressure. $\mu_a(\boldsymbol{r})$ and $\mu_s(\boldsymbol{r})$ denote the local absorption coefficient and scattering coefficient, respectively. ϕ denotes the light fluence, and $g(\boldsymbol{r})$ represents the anisotropic scattering factor. In this paper, since the focus is on the optical process of PAT imaging, we assume that the original PAT images have been accurately reconstructed [30,31], neglecting structural distortions. Additionally, in biological soft tissues, the Gruneisen coefficient varies little, so it is assumed to be constant [32]. In this case, the photoacoustic pressure $p'(\boldsymbol{r})$ in the reconstructed PAT image can be expressed as the product of the absorption coefficient μ_a and the LF distribution ϕ:
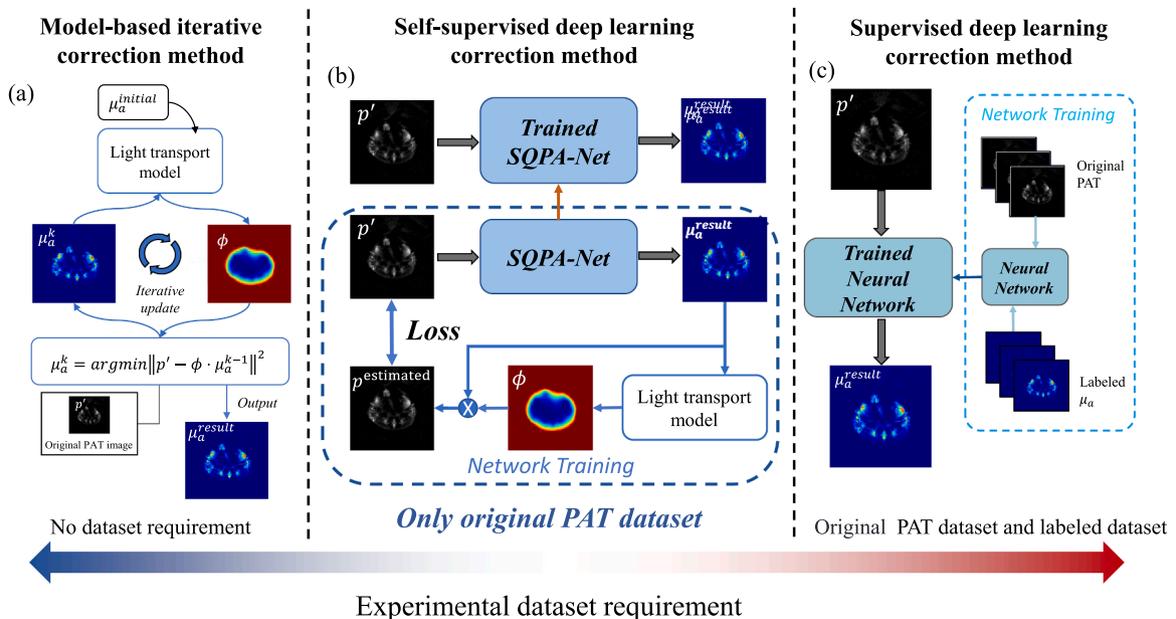
Fig. 1. The diagram of the proposed self-supervised deep learning model SQPA-Net and other methods for light fluence correction of PAT. (a) Model-based iterative correction methods. (b) Self-supervised deep learning correction method. (c) Supervised deep learning correction method.

$$p'(r) = \phi\big(\mu_a(r), \mu_s'(r)\big) \cdot \mu_a(r), \tag{2}$$

where $\mu_s'(r)$ is the reduced scattering coefficient, calculated by $\mu_s'(r) = \mu_s(r)(1 - g(r))$.

The goal of QPAT is to obtain the absorption coefficient map $\mu_a(r)$ from the reconstructed PAT image $p'(r)$. This requires the solution of the light fluence ϕ, which is dependent on $\mu_a(r)$ and $\mu_s'(r)$. Traditional LF correction method alternatively updates the unknown parameters through iterative optimization until the output of the solver matches the measured data, but its computational cost is relatively high. In comparison, recently developed DL-based methods are capable of achieving rapid LF correction, but most of them require labeled data during the training phase and overlook the benefits provided by physical information. Our work addresses these problems by proposing a self-supervised network model, SQPA-Net, and we detail its principle in following sections.

2.2. The architecture of SQPA-Net

The network architecture of the proposed SQPA-Net is illustrated in Fig. 2. The network takes the original PAT image and the reduced scattering coefficient map as input, while the output is the absorption coefficient map. The network is constrained by a physically consistent loss function to achieve self-supervised training. The two inputs of the network are first processed through two separate branches. Each branch consists of two convolutional layers and Fourier layers. The convolutional layers enhance the dimensionality of the input images, while the Fourier layers are used to extract image information. The processed data from the two branches is then combined and fed into the U-Fourier layer, which is utilized to further learn the physical information of the diffusion equation. The output from the U-Fourier layer undergoes postprocessing through two Fourier layers. Finally, after passing through two convolutional layers to reduce dimensionality, the final output is correction result.

The core of the network lies in the Fourier layer [27] and the U-Fourier layer [28], which can learn the physical information hidden within the data. The Fourier layer consists of two branches: the first branch utilizes Fourier transform to convert features into Fourier space to learn global low-frequency information, while the other branch uses convolutional layers to learn high-frequency features in the latent space. There are two reasons for learning global low information in Fourier space: first, each value in the Fourier domain is related to all values in the spatial domain of the same target, which means the network is better suited for learning global information in Fourier space. Secondly, the low-frequency information in Fourier space is more concentrated than that in the spatial domain, making it easier for the network to learn useful low-frequency information. The first branch converts the input into Fourier space using a two-dimensional Fourier transform and applies high-frequency truncation based on another hyperparameter, mode, to filter out high-frequency components and apply a linear transformation to the low-frequency components. The outputs of the two branches are summed and passed through an activation layer to the next layer. The processing procedure of the Fourier layer can be expressed as:

$$v_{t+1}(x) := \sigma\big(Wv_t(x) + \mathscr{F}^{-1}(\mathscr{F}(\kappa_\phi) \bullet \mathscr{F}(v_t)(x))\big), \tag{3}$$

The U-Fourier layer adds a compact U-Net network into the structure of the Fourier layer, aimed at enhancing the network's ability to learn high-frequency information. Its processing procedure can be expressed as:

$$v_{t+1}(x) := \sigma\big(Wv_t(x) + U_{vt}(x) + \mathscr{F}^{-1}(\mathscr{F}(\kappa_\phi) \bullet \mathscr{F}(v_t)(x))\big). \tag{4}$$

Here, σ is a nonlinear activation function, W represents a linear transformation, and U_{vt} denotes the U-Net network. κ_ϕ is the kernel function learned from the data, parameterized in Fourier space through the Fourier transform $\mathscr{F}$. The features learned in Fourier space are transformed back to the spatial domain using the inverse Fourier transform $\mathscr{F}^{-1}$.

2.3. Loss function

A key component of SQPA-Net is the self-supervised training loss function, originating from the optical processes involved in PAT imaging. Since the measured PAT image is equal to the product of the absorption coefficient map and the light fluence map, a light transport model is needed to solve for the light fluence. Here, we choose to use the diffusion equation as the light transport model. The diffusion equation is a first-order spherical harmonic expansion approximation of the radiative transfer equation, which compared to Monte Carlo numerical simulation methods, has a faster solution speed and is more for fitting into the loss function. Its expression in the frequency domain is:
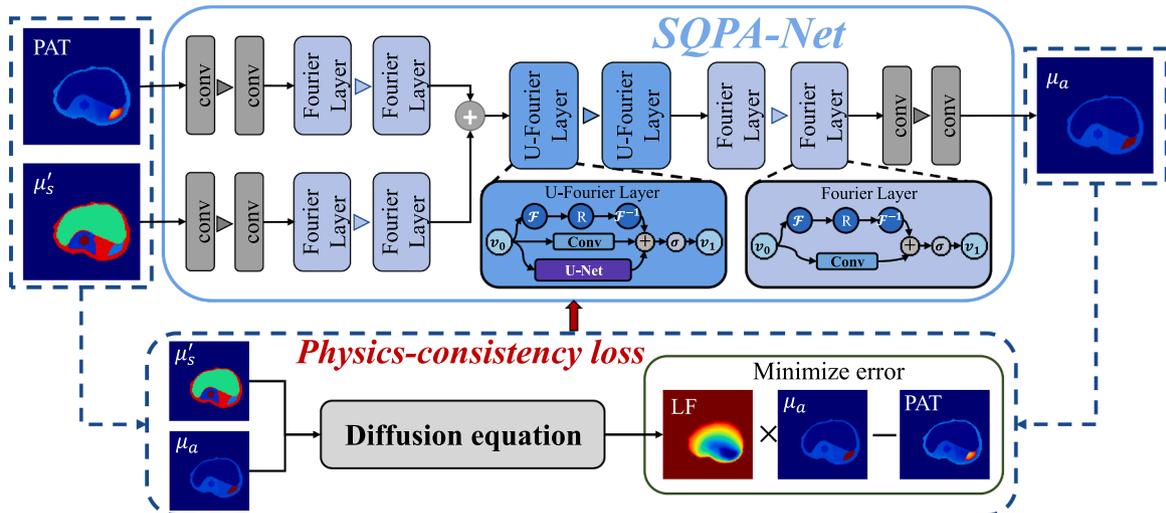
Fig. 2. The architecture of SQPA-Net for LF correction. SQPA-Net takes the original PAT image and the reduced scattering coefficient map as inputs, whereas the output is the absorption coefficient. The absorption coefficients output of SQPA-Net is used, along with the existing scattering coefficients, to solve the corresponding light fluence map based on diffusion equation. The loss function trains the network by minimizing the error between the measured PAT image and the product of the absorption coefficient and the light fluence map.

$$-\nabla \bullet D(\boldsymbol{r})\nabla\phi(\boldsymbol{r},\omega) + \left(\mu_a(\boldsymbol{r}) + \frac{i\omega}{c}\right)\phi(\boldsymbol{r},\omega) = q(\boldsymbol{r}), \qquad (5)$$

where $D = c/(3(\mu_a + \mu_s'))$ represents the light diffusion coefficient, c is the speed of light. q denotes the light source, which can be considered constant in both time and space in PAT. It can be observed that the LF distribution ϕ is jointly determined by μ_a and μ_s'. Therefore, given the distributions of μ_a and μ_s', the LF solution of Eq. (5) can be obtained using the finite element method. Since the scattering coefficient varies little in biological tissues, we generally assume it to be known [14,22, 32]. Therefore, Eq. (2) can be simplified to:

$$p'(\boldsymbol{r}) = \phi(\mu_a(\boldsymbol{r})) \cdot \mu_a(\boldsymbol{r}), \qquad (6)$$

As shown in Fig. 2, we use the original PAT image and the μ_s' image as inputs to the network, and set the network output to be the absorption coefficient map that needs to be solved, and define the loss function as

$$Loss = \sum_{\boldsymbol{r}} \left| p'(\boldsymbol{r}) - p^{estimated}(\boldsymbol{r}) \right| + \alpha(1 - SSIM(p'(\boldsymbol{r}), p^{estimated}(\boldsymbol{r}))), \qquad (7)$$

where $p^{estimated}$ represents the estimated PAT, which is the product of the network output $\mu_a^{estimated}$ and the corresponding light fluence map $\phi^{estimated}$ solved by a finite element solution of the diffusion equation [33]. Regularization parameter α is applied to the loss function to enable the appropriate balance among the two components. *SSIM* represents the structural similarity.

Compared to supervised training methods that directly minimize the error between the network output and the labels, our self-supervised method minimizes the gap between the measured PAT image and the estimated PAT image, allowing the neural network to be trained without the need for labeled data. Furthermore, since the loss function includes the diffusion equation, it enables the neural network to learn physical knowledge rather than being limited to labeled data, resulting in improved correction quality.

3. Experimental setup

3.1. PAT imaging system

The imaging equipment used is a commercial small animal multi-spectral photoacoustic tomography system (MSOT inVision128, iThera Medical, Germany). The system is equipped with a tunable laser (660–960 nm) that has a pulse width of approximately 5 ns and a repetition frequency of 10 Hz. Five pairs of laser emitters are evenly distributed at 270 degrees to provide uniform 360-degree illumination on the sample surface, creating a ring illumination approximately 8 mm in width, as present in Fig. 3(a). The ultrasound generated by the excited sample is coupled through water and transmitted to a ring-shaped array transducer composed of 128 elements, covering 270 degrees with a radius of 40.5 mm. During the imaging process, the animal is positioned

in a specialized holder that ensures alignment with the central axis of the ring-shaped transducer. The raw data are reconstructed into a 300×300 two-dimensional image using a model-based iterative image reconstruction algorithm [30]. Since the absorbed energy cannot be negative, all negative values of the reconstructed PAT maps are set to 0.

3.2. Simulation experiment

We generated absorption and reduced scattering coefficient maps using simulated organ images of healthy mice [19]. To better approximate the actual PAT imaging process, we utilized the open-source diffusion equation-based light fluence simulation software Toast++ [33] and configured the light source distribution according to the used PAT imaging system. As shown in Fig. 3(b), a circular computational grid with a radius of 40 millimeters was created in the simulation software consisting of 30 sectors and 100 rings, resulting in total of 157,291 nodes and 311,640 elements. Five light sources are evenly along the edge of the grid. The background medium for the simulation is set to pure water. The μ_a maps and μ_s' maps are placed in the center of the grid to generate the light fluence map. The simulated LF image is multiplied by the ideal μ_a image to obtain the uncorrected PAT image, to which Gaussian noise with a mean of 0 and a variance of 2×10^{-5} is further added, resulting in the final uncorrected PAT image. Following the above steps, five datasets were generated at five different illumination wavelengths, specifically 700 nm, 730 nm, 760 nm, 800 nm, and 850 nm. Each dataset contains 400 μ_a images, 400 μ_s' images, and 400 uncorrected PAT images, with 80 % allocated for the training dataset, 10 % for the validation dataset, and 10 % for the testing dataset. In the simulation experiment, it takes about 48 hours to train the SQPA-Net for each wavelength.

3.3. Animal experiment

In the animal experiments, healthy nude mice are used. All animal experiments were approved by the local Animal Ethics Committee of Southern Medical University and conducted in accordance with current guidelines. The animals were housed in ventilated cages within a temperature-controlled room, subjected to a 12-hour dark/light cycle. During the imaging process, the animals were secured on a special holder and positioned in the center of the imaging system. In order to minimize image artifacts resulting from respiratory movements, medical oxygen mixed with 1 % isoflurane (RWD Life Science Co., Ltd, China), was delivered through the breathing mask, ensuring that the respiratory rate of nude mice was maintained at 15–20 breaths per minute. Ultimately, we obtained 500 original images, and the reference absorption coefficient maps are generated using traditional iterative correction method [19], which were divided into training and testing sets in a 4:1 ratio. In later μ_a reconstruction procedure, the μ_s' is pre-determined by referring to previous literature [19], and the detail is presented in the Supplementary Material. In the animal experiment, it takes about
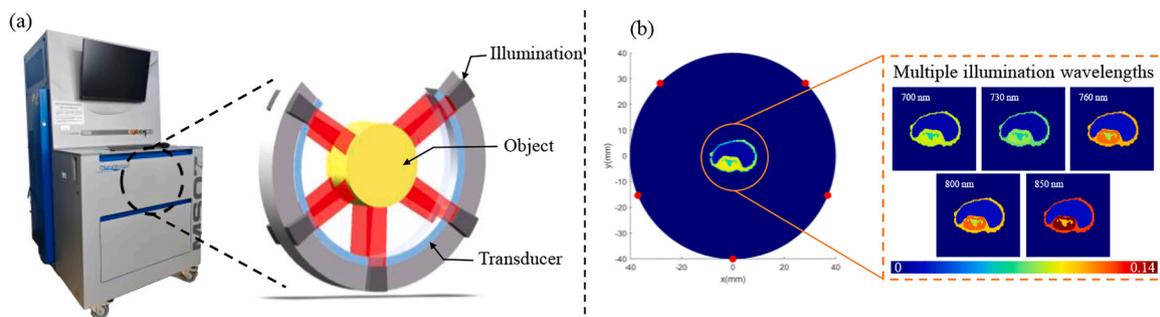
Fig. 3. (a) Schematic of the PAT system used in this study. (b) Left: Diagram of the simulation setup. Red dots represent the positions of the light sources. Right: absorption coefficient diagram of simulated mouse organ at five illumination wavelengths.

30 hours to train the SQPA-Net for each wavelength.

3.4. Evaluation metrics

For all experiment results, we use root mean square error (RMSE), structural similarity (SSIM), and peak signal-to-noise ratio (PSNR) to evaluate the quality of corrected images, and their expressions are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(T-M)^2}, \tag{8}$$

$$PSNR = 20\log_{10}\left(\frac{MAX}{RMSE}\right), \tag{9}$$

$$SSIM = \frac{(2\mu_M\mu_T + C_1)(2\sigma_{MT} + C_2)}{(\mu_M^2 + \mu_T^2 + C_1)(\sigma_M^2 + \sigma_T^2 + C_2)}, \tag{10}$$

where T represents the ground truth, M represents the result of the experiment, n represents number of pixels. MAX is the max value of the ground truth. μ_M and μ_T represent the mean of M and T, respectively, σ_M^2 and σ_T^2 are the corresponding variance. σ_{MT} is the cross-covariance for M and T sequentially. The C_1 and C_2 are given by:

$$C_1 = (0.01 \times MAX)^2, C_2 = (0.03 \times MAX)^2. \tag{11}$$

3.5. Deep learning setup

The algorithm development platform used in this study is a desktop computer based on the AMD Ryzen 5–5600 CPU and NVIDIA RTX 3060 GPU. For simulation experiments, the self-supervised neural network is trained using the Adam optimizer in 80 epochs, with an initial learning rate of 0.002. The learning rate decreasing to 0.1 of its original value every 20 epochs. For animal experiments, the self-supervised neural network is trained using the Adam optimizer in 40 epochs, with an initial learning rate of 0.002. The learning rate decreasing to 0.1 of its original value every 10 epochs. We set regularization parameter α of loss function to 100. Detailed descriptions of the comparative methods can be found in the Supplementary Material.

4. Results

4.1. Simulation results

In the simulation experiments, we compared four different correction methods: traditional iterative correction (TIC) [19], SQPA-Net, U-Net [34], and Dual-path Network [22]. The visual comparison of the experimental results at 850 nm illumination wavelength is shown in Fig. 4. It is evident from Fig. 4(a) and Fig. 4(b) that due to the influence of the light fluence, the energy of peripheral tissues in the original PAT image is excessively high compared to the ideal absorption coefficient diagram (as indicated by the blue arrows). Meanwhile, due to the attenuation of the light fluence, the signal of deeper tissues is too low (as
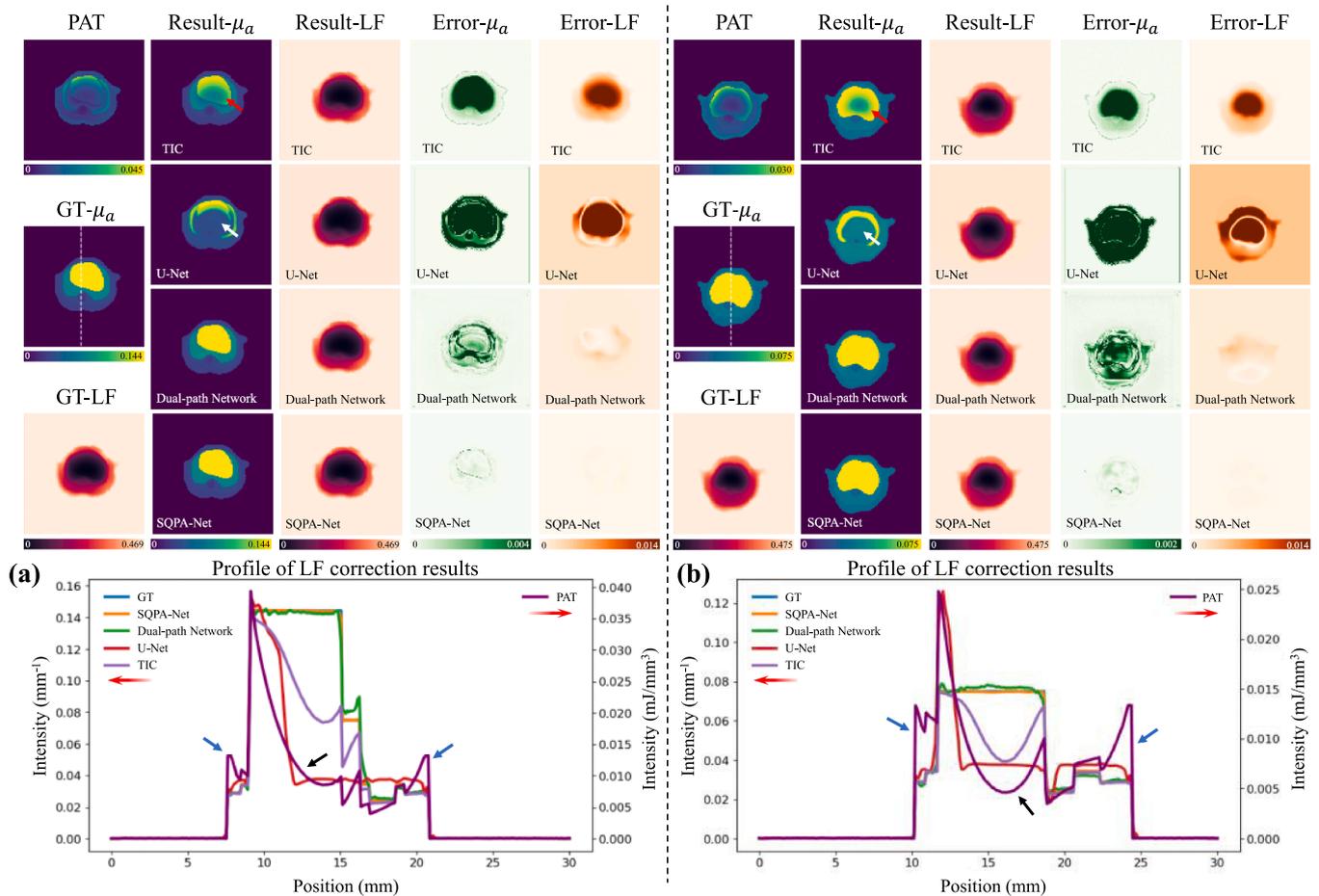
Fig. 4. Visual comparison of LF correction results of different methods at 850 nm illumination wavelength in simulation experiment. PAT: un-corrected PAT maps. GT-μ_a: ideal μ_a maps. GT-LF: ideal LF maps. Result-μ_a: the correction results of μ_a from different methods. Result-LF: the resulting LF from different methods. Error-μ_a: absolute error between the correction result of μ_a and ideal μ_a map. Error-LF: absolute error between the obtained LF and ideal LF map. TIC: traditional iterative correction method. (a)(b): Profiles of PAT and μ_a maps drawn along the white dashed line in GT-μ_a.

indicated by the black arrow). Meanwhile, due to the attenuation of the light fluence, the signal of the deeper tissues is too low (as indicated by the black arrow). Therefore, the original PAT image cannot accurately reflect the proportion of absorption coefficients in various parts of the tissue. Furthermore, all four correction methods can perform a certain degree of LF correction, but the quality of correction varies. TIC reconstructs the absorption coefficient maps for most tissues accurately. However, it cannot fully correct for large areas of high absorption coefficient organs within the tissues (as indicated by the red arrow), which demonstrates that the traditional iterative method has certain limitations in correcting for deep high absorption coefficients. Correspondingly, the LF results of TIC also exhibited significant errors in the deeper regions of the tissue. The correction results of the SQPA-Net are the closest to the ideal absorption coefficient maps among the four correction methods, including the complete reconstruction of regions with high internal absorption coefficients, with well-defined organ contours. As seen in Fig. 4(a) and Fig. 4(b), SQPA-Net successfully performs LF correction at different tissue combinations and different depths. The light fluence maps produced by SQPA-Net are also the closest to the ideal LF maps among the four methods. Nevertheless, there are still minor discrepancies as observed in the error map. The correction results of U-Net are the lowest quality among four LF correction methods. The experimental results show that the regions of high absorption coefficients in deep tissues are missing (as indicated by the white arrow), and the edges of different tissues are blurred. The error map reveals that the results of U-Net deviate significantly from the ideal image. Compared to U-Net, the correction results of the Dual-path Network are significantly improved. It can reconstruct regions of high absorption coefficients within the tissues while clearly delineating the edges of different tissues. The error map indicates that there are fewer errors compared to U-Net, but the errors of Dual-path Network are present throughout the tissues.

Fig. 5 presents the visual comparison of LF correction results of different methods at three different illumination wavelengths. Similar to the results at different locations, SQPA-Net produced the best correction results under various illumination wavelengths, with the least error. TIC still fails to achieve accurate correction in areas with high internal absorption coefficients. The other two U-Net-based methods produced significant error, leading to unsatisfactory results.

To comprehensively compare the correction results of the four methods, we conducted experiments on all four methods at five different illumination wavelengths (700 nm, 730 nm, 760 nm, 800 nm, 850 nm), and evaluated the correction results using quantitative metrics. Table 1 lists the quantitative results of the four LF correction methods. At five illumination wavelengths, TIC reaches a RMSE of 0.00088 ± 0.00204, a PSNR of 50.762 ± 12.41 and a 1-SSIM of 0.0034 ± 0.0025. SQPA-Net reaches a RMSE of 0.00014 ± 0.0002, a PSNR of 54.42 ± 6.57 and a 1-SSIM of 0.0020 ± 0.0022. U-Net reaches a RMSE of 0.00290 ± 0.00368, a PSNR of 27.52 ± 5.78 and a 1-SSIM of 0.0365 ± 0.0297. Dual-path Network reaches a RMSE of 0.00089 ± 0.00101, a PSNR of 36.78 ± 6.19 and a 1-SSIM of 0.0418 ± 0.0589. From the quantitative metrics, it can be observed that SQPA-Net achieved the best calibration results across five illumination wavelengths, followed by TIC, with Dual-path Network ranking third, and U-Net performing the worst in calibration, which is consistent with the visual comparison results. Additionally, in Table 1, we also reported the average time consumed for processing a single PAT map of four LF correction methods. To process a single image at five illumination wavelengths, TIC needs 83.47 s $\pm$ 3.54 s, SQPA-Net needs 0.03 s $\pm$ 0.01 s, U-Net needs 0.01 s $\pm$ 0.00 s, Dual-path needs 0.01 s $\pm$ 0.00 s. Compared to traditional iterative methods, SQPA-Net has achieved a significant improvement in correction speed.

To further evaluate the effectiveness and precision of four distinct LF correction methods, we employed the linear spectral unmixing method [35] to calculate the distributions of oxyhemoglobin (HbO2), deoxyhemoglobin (Hb), and oxygen saturation (sO2) based on the experimental results obtained. Fig. 6 shows a visual comparison of linear spectral unmixing results at two different positions. Since the high internal absorption coefficient areas cannot be accurately reconstructed, the spectral unmixing results of TIC exhibit errors in the same areas. In contrast, the results of SQPA-Net are the closest to the ideal image among the four LF correction methods. The spectral separation results of U-Net and Dual-path Network are both unsatisfactory and deviate significantly from the ideal image. Table 2 lists the quantification metrics of sO2 from the four methods. Consistent with the visual comparison results, SQPA-Net achieves the best quantitative metrics, with a RMSE of 0.0035 ± 0.0043, a PSNR of 51.765 ± 8.439, and a 1-SSIM of 0.0018 ± 0.0045. Compared to TIC, the RMSE of SQPA-Net decreased by 0.0017, the PSNR increased by 2.572, and the 1-SSIM decreased by 0.0029.
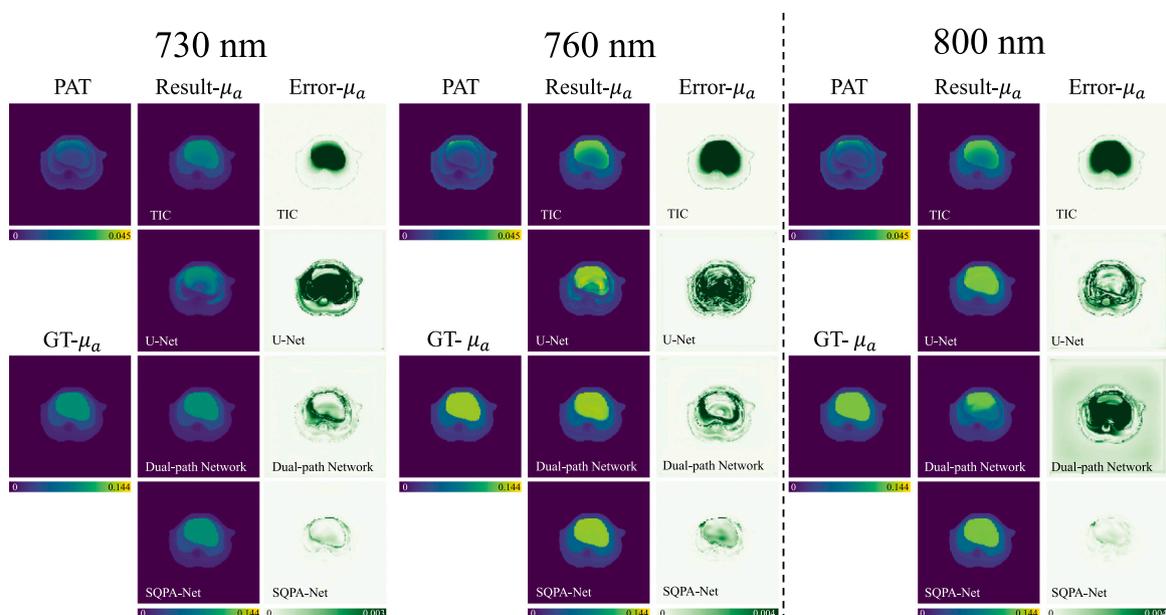
Fig. 5. LF correction results at three different illumination wavelengths in simulation experiment. PAT: un-corrected PAT maps. GT-μ_a: ideal μ_a maps. Result-μ_a: the correction results of μ_a from different methods. Error-μ_a: absolute error between the correction result of μ_a and ideal μ_a map.

Table 1

Quantitative results (mean with standard deviation) of different LF correction methods on simulation data of five different illumination wavelengths.

Wavelength (nm)	Method	RMSE	PSNR	1-SSIM	Time (s)
700	TIC	0.00054 ± 0.00125	50.66 ± 10.46	0.0042 ± 0.0027	84.78 ± 0.35
	U-Net	0.00198 ± 0.00181	27.79 ± 5.24	0.0474 ± 0.0350	**0.01 ± 0.00**
	Dual-path Network	0.00078 ± 0.00022	33.60 ± 5.06	0.0915 ± 0.0633	**0.01 ± 0.00**
	SQPA-Net	**0.00013 ± 0.00016**	**54.64 ± 7.18**	**0.0028 ± 0.0028**	0.03 ± 0.01
730	TIC	0.00026 ± 0.00059	51.97 ± 8.23	0.0047 ± 0.0034	82.09 ± 0.39
	U-Net	0.00226 ± 0.00214	26.44 ± 5.47	0.0339 ± 0.0270	**0.01 ± 0.00**
	Dual-path Network	0.00027 ± 0.00018	43.05 ± 3.29	**0.0030 ± 0.0013**	**0.01 ± 0.00**
	SQPA-Net	**0.00011 ± 0.00013**	**53.75 ± 5.41**	0.0034 ± 0.0028	0.03 ± 0.01
760	TIC	0.00109 ± 0.00225	50.11 ± 13.55	0.0029 ± 0.0018	84.76 ± 0.31
	U-Net	0.00282 ± 0.00243	27.50 ± 4.84	0.0473 ± 0.0342	**0.01 ± 0.00**
	Dual-path Network	0.00068 ± 0.00041	38.05 ± 4.07	0.0093 ± 0.0042	**0.01 ± 0.00**
	SQPA-Net	**0.00018 ± 0.0002**	**51.64 ± 5.89**	**0.0018 ± 0.0013**	0.03 ± 0.00
800	TIC	0.00090 ± 0.00179	50.62 ± 13.03	0.0027 ± 0.0015	79.50 ± 5.78
	U-Net	0.00121 ± 0.00070	33.16 ± 2.68	0.0123 ± 0.0043	**0.01 ± 0.00**
	Dual-path Network	0.00195 ± 0.00172	30.10 ± 4.83	0.0962 ± 0.0651	**0.01 ± 0.00**
	SQPA-Net	**0.00009 ± 0.00010**	**57.13 ± 6.39**	**0.0010 ± 0.0010**	0.03 ± 0.00
850	TIC	0.00163 ± 0.00307	50.46 ± 15.41	0.0025 ± 0.0020	86.23 ± 0.32
	U-Net	0.00624 ± 0.00617	22.73 ± 4.83	0.0414 ± 0.0205	**0.01 ± 0.00**
	Dual-path Network	0.00077 ± 0.00052	39.10 ± 3.78	0.0089 ± 0.0038	**0.01 ± 0.00**
	SQPA-Net	**0.00018 ± 0.00033**	**54.94 ± 6.61**	**0.0008 ± 0.0006**	0.03 ± 0.00

4.2. Animal imaging results

To further verify the performance of SQPA-Net, we conducted LF correction experiments on animal PAT images. Since it is difficult to obtain the ideal μ_a maps and LF maps of animals, we used the correction results of traditional iterative reconstruction methods as a reference to evaluate three deep learning-based correction methods. Fig. 7 presents a visual comparison of the correction results at two different positions at 700 nm illumination wavelength. Firstly, all three correction methods are capable of performing LF correction, but their effectiveness varies. From the absolute error maps, it can be seen that the U-Net correction results have the most errors among three correction methods. Especially in areas with high absorption coefficients (as indicated by the yellow arrows), the error of U-Net increases. In contrast, the errors of the other two correction methods are much smaller, but they also show errors in areas with high absorption coefficients. The SQPA-Net has lower errors in high absorption coefficient areas compared to the Dual-path Network (as indicated by the red arrows). Additionally, at two different positions, the SQPA-Net maintained the smallest errors among the three methods, demonstrating its consistent correction ability at different positions.

Additionally, we conducted experiments at five different illumination wavelengths (700 nm, 730 nm, 760 nm, 800 nm, 850 nm) to verify the performance of three different correction methods under various illumination wavelengths. Fig. 8 presents a visual comparison of correction results at three different illumination wavelengths (760 nm, 800 nm, 850 nm). From the figure, it can be seen that the U-Net method exhibits the largest error among the three correction methods at three illumination wavelengths, with errors increasing in regions with high absorption coefficients (as indicated by the yellow arrows). This is consistent with experimental results of different positions, indicating that the performance of U-Net is not ideal. Compared to U-Net, Dual-path Network exhibits fewer errors. However, the error map shows that at illumination wavelength of 800 nm, the Dual-path Network has larger errors compared to illumination wavelengths of 760 nm and 850 nm. In contrast, SQPA-Net maintains consistent errors at three different illumination wavelengths and has the least errors among three correction methods. This demonstrates that the LF correction capability of SQPA-Net is unaffected by different illumination wavelengths.

In order to intuitively compare the correction effects of different methods under five illumination wavelengths, we conducted a quantitative evaluation of the correction results. Table 3 shows the quantitative metrics of the three correction methods across three illumination wavelengths. The average RMSE of SQPA-Net at five illumination wavelengths reaches 0.00031 ± 0.00032, average PSNR reaches 53.42

± 3.80 and average 1-SSIM reaches 0.0005 ± 0.0007. The average RMSE of U-Net reaches 0.00137 ± 0.00087, average PSNR reaches 39.45 ± 3.73 and average 1-SSIM reaches 0.0068 ± 0.0058. The average RMSE of Dual-path Network reaches 0.00063 ± 0.00060, average PSNR reaches 47.29 ± 4.37 and average 1-SSIM reaches 0.0016 ± 0.0016. Among the three correction methods, SQPA-Net has the best quantitative metrics, Dual-path Network ranks second, and U-Net performs the worst, which is consistent with the visual comparison results. Moreover, the quantitative metric PSNR indicates that the correction quality of SQPA-Net is relatively consistent at five different illumination wavelengths, with PSNR values all above 50 dB, with a difference of 3.63 dB between the highest and lowest values. By contrast, the PSNR of the Dual-path Network fluctuates significantly, with a difference of 6.97 dB between the highest and lowest values, indicating that the wavelength significantly affects its correction quality.

We also recorded the average time of different correction methods to process a single image to evaluate the correction efficiency of each method. Table 4 shows the correction time for the four different methods at five illumination wavelengths. The average computation time of TIC, SQPA-Net, U-Net, and Dual-path Network for a single image is 106.31 s ± 3.18 s, 0.02 s ± 0.00 s, 0.01 s ± 0.00 s and 0.01 s ± 0.00 s, respectively. Deep-learning-based correction methods have improved speed by thousands of times compared to traditional iterative methods, resulting in a significant enhancement in correction efficiency.

5. Discussions

Based on the results of both simulation and animal experiments, it can be concluded that the proposed self-supervised network can achieve a highly efficient light fluence correction process for QPAT. Through visual comparison and quantitative evaluation, SQPA-Net demonstrates correction performance that matches or even exceeds that of traditional iterative correction methods. Compared to traditional iterative methods that take tens of seconds to process a single image, SQPA-Net improves correction speed by over thousand times at only 0.02–0.03 seconds, thus making it a potential approach for real-time QPAT imaging. Compared to state-of-the-art DL-based methods, due to the utilization of a self-supervised network training scheme, SQPA-Net eliminates the need for high-quality labeled datasets. This significantly reduces the cost and time resources needed for dataset preparation in PAT imaging, where obtaining accurate animal absorption coefficient maps and light fluence maps is challenging.
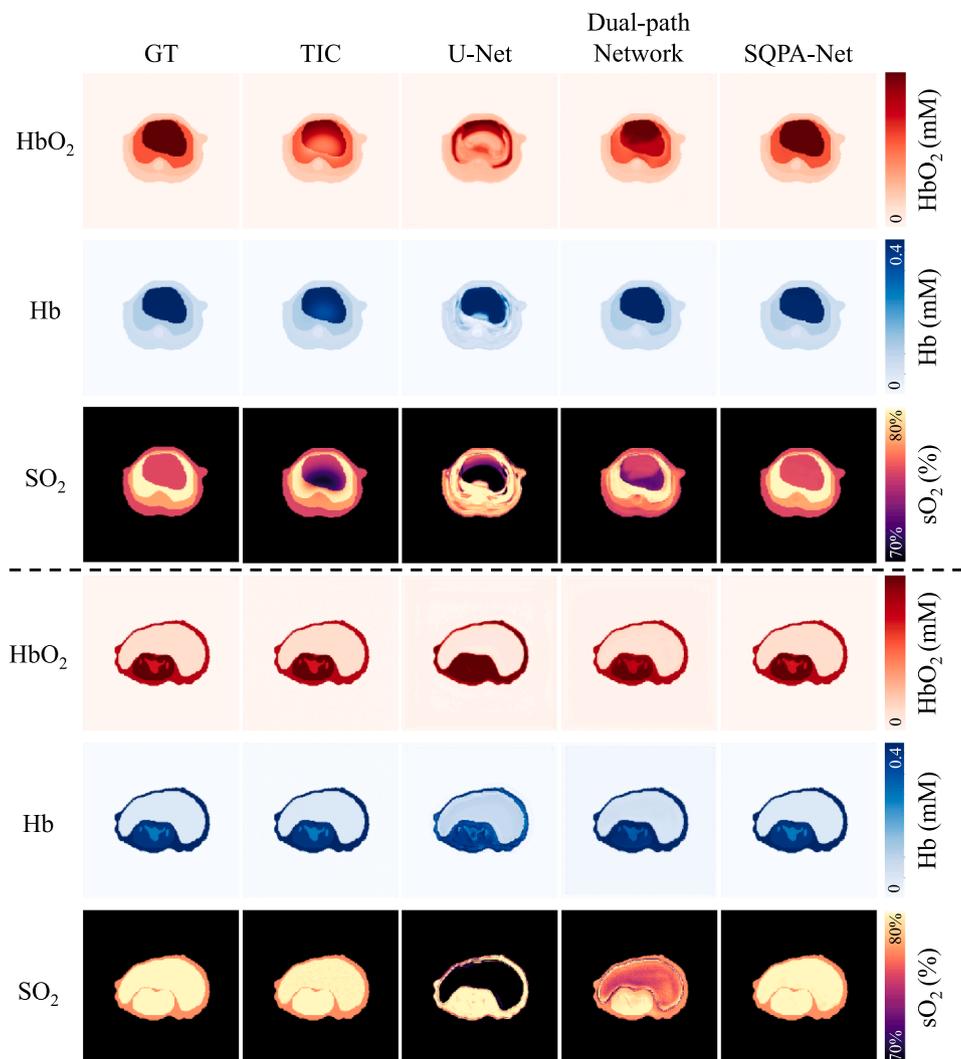
Fig. 6. The results of spectral un-mixing from four different LF correction methods. GT: ideal map. TIC: traditional iterative correction method.

Table 2
Quantitative results (mean with standard deviation) of sO$_2$ from four different correction methods.

	TIC	U-Net	Dual-path Network	SQPA-Net
RMSE	0.0052 ± 0.0060	0.0387 ± 0.0179	0.0176 ± 0.0081	$\mathbf{0.0035 \pm 0.0043}$
PSNR	49.193 ± 9.993	27.43 ± 4.405	34.503 ± 5.273	$\mathbf{51.765 \pm 8.439}$
1-SSIM	0.0047 ± 0.0052	0.0564 ± 0.0366	0.0171 ± 0.0151	$\mathbf{0.0018 \pm 0.0045}$

As can be found in the simulation experiment, traditional iterative correction methods cannot fully restore the absorption coefficients in large areas with high absorption coefficients, where rapid attenuation of light fluence results in low image intensity. In comparison, our SQPA-Net successfully restores the absorption coefficients in deep tissue regions with high absorption, and further enables improved spectral unmixing accuracy. This finding indicate that our SQPA-Net may achieve better LF correction performance than traditional iterative correction methods in real-world PAT imaging. However, this cannot be verified in our current experimental setting because the real-world μ_a map is not available.

Furthermore, comparing to current DL-based methods, one feature of our SQPA-Net is that it maintains consistent correction capability when dealing with various illumination wavelengths and imaging positions. This can be found in the simulation experiments where SQPA-Net maintains a consistent error level across two different positions while the Dual-path Network exhibits significantly larger errors. The same finding can be confirmed in the imaging experiments of small animal. Moreover, SQPA-Net also maintains consistent error across different illumination wavelengths, where the difference between highest and lowest average PSNR of SQPA-Net is 5.49 dB (simulation) and 3.63 dB (animal), while that of Dual-path Network is 12.95 dB (simulation) and 6.97 dB (animal). In addition, we tested the performance of SQPA-Net in different illumination and noisy environments, and the experimental results can be found in the supplementary materials. The results further show that our SQPA-Net maintains high-quality results even under
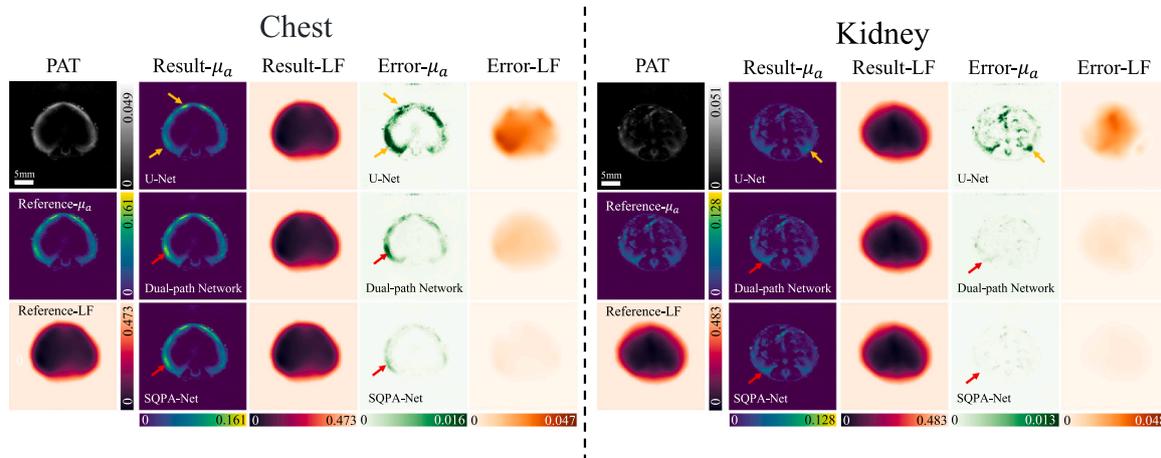
Fig. 7. LF correction results at two different positions in small animal imaging experiment. Reference-μ_a: correction results of μ_a from traditional iterative correction method. Reference-LF: correction results of LF from traditional iterative correction method. PAT: uncorrected PAT image. Result-μ_a: correction results of μ_a from different methods. Result-LF: the resulting LF from different methods. Error-μ_a: absolute error between the correction result of μ_a and reference μ_a map. Error-LF: absolute error between the obtained LF and reference LF map.
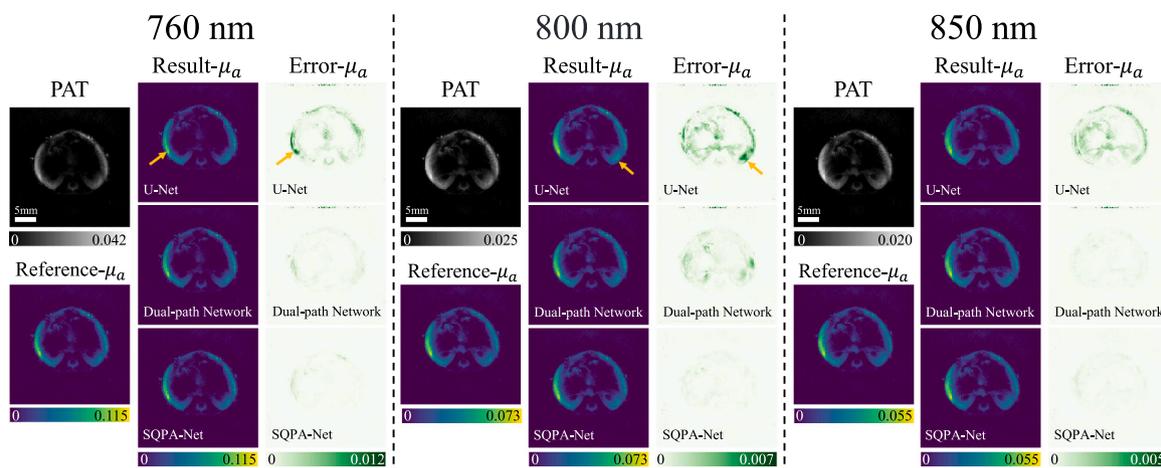
Fig. 8. LF correction results at three different illumination wavelengths in small animal imaging experiment. Reference-μ_a: correction results of μ_a from traditional iterative correction method. PAT: uncorrected PAT image. Result-μ_a: LF correction results of different methods. Error-μ_a: absolute error between the correction result of μ_a and ideal μ_a map.

conditions of reduced illumination and increased noise.

Additionally, the incorporation of physical information is also crucial for SQPA-Net to achieve self-supervised light fluence correction. During the training process, we use a gradient descent learning rate to ensure network convergence and achieve high-quality light fluence correction. In order to test the effectiveness of physical information, we train SQPA-Net using a fully supervised method without changing the type of loss function and compare it with the self-supervised SQPA-Net (experimental results can be found in the supplementary materials). The quantitative results indicate that the correction quality of the self-supervised SQPA-Net is superior to that of the fully supervised SQPA-Net, demonstrating that physical information can improve the performance of our network.

Despite these advantages, the method proposed has some limitations which can be further improved in future works. Firstly, due to the difficulty of solving the absorption coefficient map and the scattering coefficient map simultaneously [32], we assume that the reduced scattering coefficient is known in our current model. Due to the difficulty in distinguishing the boundaries of different tissues in the original

PAT images, we have applied a uniform reduced scattering coefficient to the image target. Therefore, there is a certain degree of error between our correction results and the actual situation. We have attempted to remove μ_s' from the network input, yet the performance of SQPA-Net with only PAT as input deteriorated significantly (experimental results can be found in the supplementary materials). How to solve for both the absorption coefficient and the scattering coefficient needs future research. Furthermore, in practical PAT imaging, the LF varies across three-dimensional space. The LF distribution is affected not only by tissue within the same slice but also by that between different slices. In our current method, we only examine the LF in two-dimension, which may result in discrepancies in the generated LF map. Finally, our experiments are conducted under the assumption that the original PAT images are perfectly reconstructed, yet image reconstruction significantly affects LF estimation [36]. Acoustic heterogeneity of animal tissues poses challenges in achieving perfect reconstruction, and artifacts or noise in the reconstructed images may be misinterpreted as optical absorbers after LF correction. The effect of image reconstruction on LF correction needs to be investigated in future works.

Table 3

Quantified correction results (mean with standard deviation) of animal experiments at five illumination wavelengths.

Wavelength (nm)	Method	RMSE	PSNR	1-SSIM
700	U-Net	0.00214 ± 0.00112	37.07 ± 2.96	0.0099 ± 0.0065
	Dual-path Network	0.00102 ± 0.00079	44.08 ± 3.75	0.0033 ± 0.0020
	SQPA-Net	**0.00050 ± 0.00051**	**51.29 ± 3.71**	**0.0007 ± 0.0010**
730	U-Net	0.00116 ± 0.00040	39.86 ± 3.00	0.0062 ± 0.0036
	Dual-path Network	0.00052 ± 0.00048	48.54 ± 3.76	0.0008 ± 0.0007
	SQPA-Net	**0.00025 ± 0.00024**	**54.92 ± 3.67**	**0.0003 ± 0.0005**
760	U-Net	0.00088 ± 0.00039	43.07 ± 2.91	0.0025 ± 0.0017
	Dual-path Network	0.00058 ± 0.00046	47.47 ± 3.70	0.0012 ± 0.0009
	SQPA-Net	**0.00026 ± 0.00021**	**54.57 ± 3.55**	**0.0003 ± 0.0004**
800	U-Net	0.00155 ± 0.00091	37.67 ± 3.18	0.0088 ± 0.0069
	Dual-path Network	0.00068 ± 0.00061	45.33 ± 3.27	0.0020 ± 0.0014
	SQPA-Net	**0.00032 ± 0.00028**	**52.37 ± 3.55**	**0.0007 ± 0.0007**
850	U-Net	0.00112 ± 0.00060	39.57 ± 3.34	0.0068 ± 0.0056
	Dual-path Network	0.00035 ± 0.00036	51.05 ± 3.60	0.0006 ± 0.0006
	SQPA-Net	**0.00023 ± 0.00019**	**53.95 ± 3.19**	**0.0005 ± 0.0005**

Table 4

The LF correction time (s) of methods in animal experiments with five illumination wavelengths.

	700 nm	730 nm	760 nm	800 nm	850 nm
TIC	106.74 ± 3.01	105.00 ± 0.43	105.78 ± 1.09	107.38 ± 4.13	106.74 ± 3.01
U-Net	**0.01 ± 0.00**	**0.01 ± 0.00**	**0.01 ± 0.00**	**0.01 ± 0.00**	**0.01 ± 0.00**
Dual-path Network	**0.01 ± 0.00**	**0.01 ± 0.00**	**0.01 ± 0.00**	**0.01 ± 0.00**	**0.01 ± 0.00**
SQPA-Net	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00

6. Conclusion

In this study, we proposed SQPA-Net, a self-supervised quantitative photoacoustic neural network that incorporates diffusion equation-based light transport model for LF correction in PAT imaging. Our method successfully achieves self-supervised network training, and thus effectively reduces the cost of preparing labeled datasets while significantly improving LF correction speed. We tested the proposed method on simulation and small animal imaging experiments, and the results demonstrate the effectiveness and efficiency of our method.

Funding

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.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.pacs.2025.100684.

Data availability

Data will be made available on request.

References

[1] L.V. Wang, J. Yao, A practical guide to photoacoustic tomography in the life sciences, Nat. Methods 13 (8) (2016) 627–638.

[2] L.V. Wang, S. Hu, Photoacoustic tomography: in vivo imaging from organelles to organs, Science 335 (6075) (2012) 1458–1462.

[3] S. Zhang, L. Qi, X. Li, J. Liu, W. Chen, Photoacoustic imaging of living mice enhanced with a low-cost contrast agent, Biomed. Opt. Express 10 (11) (2019) 5744.

[4] 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.

[5] S. Tzoumas, N. Deliolanis, S. Morscher, V. Ntziachristos, Unmixing Molecular Agents From Absorbing Tissue in Multispectral Optoacoustic Tomography, IEEE Trans. Med. IMAGING MI 33 (1) (2014) 48–60.

[6] Y. Zhong, X. Zhang, Z. Mo, S. Zhang, L. Nie, W. Chen, L. Qi, Spiral scanning and self-supervised image reconstruction enable ultra-sparse sampling multispectral photoacoustic tomography, Photoacoustics 39 (2024) 100641.

[7] S. Zhang, L. Qi, X. Li, Z. Liang, X. Sun, J. Liu, L. Lu, Y. Feng, W. Chen, MRI Information-Based Correction and Restoration of Photoacoustic Tomography, Ieee T Med Imaging 41 (9) (2022) 2543–2555.

[8] 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.

[9] S. Zhang, L. Qi, X. Li, Z. Liang, J. Wu, S. Huang, J. Liu, Z. Zhuang, Y. Feng, Q. Feng, In vivo co-registered hybrid-contrast imaging by successive photoacoustic tomography and magnetic resonance imaging, Photoacoustics 31 (2021).

[10] K. Huda, D.J. Lawrence, W. Thompson, S.H. Lindsey, C.L. Bayer, In vivo noninvasive systemic myography of acute systemic vasoactivity in female pregnant mice, Nat. Commun. 14 (1) (2023) 6286.

[11] G. Diot, S. Metz, A. Noske, E. Liapis, B. Schroeder, S.V. Ovsepian, R. Meier, E. Rummeny, V. Ntziachristos, Multispectral Optoacoustic Tomography (MSOT) of Human Breast Cancer, Clin. Cancer Res 23 (22) (2017) 6912–6922.

[12] Y. Chen, H. Yang, Y. Luo, Y. Niu, M. Yu, S. Deng, X. Wang, H. Deng, H. Chen, L. Gao, Photoacoustic Tomography with Temporal Encoding Reconstruction (PATTERN) for cross-modal individual analysis of the whole brain, Nat. Commun. 15 (1) (2024).

[13] B. Cox, J.G. Laufer, S.R. Arridge, P.C. Beard, Quantitative spectroscopic photoacoustic imaging: a review, J. Biomed. Opt. 17 (6) (2012) 061202.

[14] B.T. Cox, S.R. Arridge, K.P. Köstli, P.C. Beard, Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method, Appl. Opt. 45 (8) (2006) 1866–1875.

[15] L. Ulrich, L. Ahnen, H.G. Akarçay, S.S. Majos, M. Jaeger, K.G. Held, M. Wolf, M. Frenz, Spectral correction for handheld optoacoustic imaging by means of near-infrared optical tomography in reflection mode, J. Biophotonics 12 (1) (2019) e201800112.

[16] K. Daoudi, A. Hussain, E. Hondebrink, W. Steenbergen, Correcting photoacoustic signals for fluence variations using acousto-optic modulation, Opt. Express 20 (13) (2012) 14117–14129.

[17] Z. Wang, W. Tao, H. Zhao, The Optical Inverse Problem in Quantitative Photoacoustic Tomography: A Review, Photonics 10 (5) (2023) 487.

[18] Y. Liu, H. Jiang, Z. Yuan, Two schemes for quantitative photoacoustic tomography based on Monte Carlo simulation, Med Phys. 43 (7) (2016) 3987.

[19] 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.

[20] C. Cai, K. Deng, C. Ma, J. Luo, End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging, Opt. Lett. 43 (12) (2018) 2752–2755.
[21] G.P. Luke, K. Hoffer-Hawlik, A.C. Van Namen, R. Shang, O-Net: A Convolutional Neural Network for Quantitative Photoacoustic Image Segmentation and Oximetry, arXiv e-prints (2019).
[22] J. Li, C. Wang, T. Chen, T. Lu, S. Li, B. Sun, F. Gao, V. Ntziachristos, Deep learning-based quantitative optoacoustic tomography of deep tissues in the absence of labeled experimental data, Optica 9 (1) (2022) 32–41.
[23] Z. Liang, S. Zhang, Z. Liang, Z. Mo, X. Zhang, Y. Zhong, W. Chen, L. Qi, Deep learning acceleration of iterative model-based light fluence correction for photoacoustic tomography, Photoacoustics 37 (2024) 100601.
[24] J.H. Nölke, T.J. Adler, M. Schellenberg, K.K. Dreher, N. Holzwarth, C.J. Bender, M. D. Tizabi, A. Seitel, L. Maier-Hein, Photoacoustic Quantification of Tissue Oxygenation Using Conditional Invertible Neural Networks, IEEE Trans. Med. Imaging 43 (9) (2024) 3366–3376.
[25] Z. Wang, W. Tao, Z. Zhang, H. Zhao, Extractor-attention-predictor network for quantitative photoacoustic tomography, Photoacoustics 38 (2024) 100609.
[26] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707.
[27] Z.Y. Li, N.B. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A.M. Stuart, A. Anandkumar, Fourier Neural Operator for Parametric Partial Differential Equations, ArXiv abs/2010.08895 (2020).
[28] G. Wen, Z. Li, K. Azizzadenesheli, A. Anandkumar, S.M. Benson, U-FNO—An enhanced Fourier neural operator-based deep-learning model for multiphase flow, Adv. Water Resour. 163 (2022) 104180.
[29] L. Huang, H. Chen, T. Liu, A. Ozcan, Self-supervised learning of hologram reconstruction using physics consistency, Nat. Mach. Intell. 5 (8) (2023) 895–907.
[30] 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.
[31] 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.
[32] F.M. Brochu, J. Brunker, J. Joseph, M.R. Tomaszewski, S. Morscher, S.E. Bohndiek, Towards Quantitative Evaluation of Tissue Absorption Coefficients Using Light Fluence Correction in Optoacoustic Tomography, IEEE Trans. Med Imaging 36 (1) (2016) 322–331.
[33] M. Schweiger, S. Arridge, The Toast++ software suite for forward and inverse modeling in optical tomography, J. Biomed. Opt. 19 (4) (2014) 040801.
[34] T. Chen, T. Lu, S. Song, S. Miao, F. Gao, J. Li, A deep learning method based on U-Net for quantitative photoacoustic imaging, SPIE (2020).
[35] 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.
[36] R. Zhang, Rt A'Dawiah, T.W.J. Choo, X. Li, G. Balasundaram, Y. Qi, Y. Goh, R. Bi, M. Olivo, Navigating challenges and solutions in quantitative photoacoustic imaging, Appl. Phys. Rev. 11 (3) (2024) 031308.

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.

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.

Long Chen received his bachelor's degree from Southern Medical University in 2024. He is currently pursuing the master's degree in the school of Biomedical Engineering, Southern Medical University. His research interests include image processing and photoacoustic tomography.

Danni Wang is pursuing the master's degree in the school of Biomedical Engineering, Southern Medical University. Her research interests include photoacoustic tomography and photoacoustic imaging system for oral disease.

Chaobin Hu received his bachelor's degree from Southern Medical University in 2023, He is currently pursuing the master's degree in the school of Biomedical Engineering, Southern Medical University. His research focuses on deep learning-based photoacoustic tomography image reconstruction.

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.