

Back-propagation neural network-based reconstruction algorithm for diffuse optical tomography

Jinchao Feng
Qiuwan Sun
Zhe Li
Zhonghua Sun
Kebin Jia

Back-propagation neural network-based reconstruction algorithm for diffuse optical tomography

Jinchao Feng,^{a,b} Qiuwan Sun,^a Zhe Li,^{a,b,*} Zhonghua Sun,^{a,b} and Kebin Jia^{a,b,*}

^aBeijing University of Technology, Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing, China

^bBeijing Laboratory of Advanced Information Networks, Beijing, China

Abstract. Diffuse optical tomography (DOT) is a promising noninvasive imaging modality and is capable of providing functional characteristics of biological tissue by quantifying optical parameters. The DOT image reconstruction is ill-posed and ill-conditioned, due to the highly diffusive nature of light propagation in biological tissues and limited boundary measurements. The widely used regularization technique for DOT image reconstruction is Tikhonov regularization, which tends to yield oversmoothed and low-quality images containing severe artifacts. It is necessary to accurately choose a regularization parameter for Tikhonov regularization. To overcome these limitations, we develop a noniterative reconstruction method, whereby optical properties are recovered based on a back-propagation neural network (BPNN). We train the parameters of BPNN before DOT image reconstruction based on a set of training data. DOT image reconstruction is achieved by implementing a single evaluation of the trained network. To demonstrate the performance of the proposed algorithm, we compare with the conventional Tikhonov regularization-based reconstruction method. The experimental results demonstrate that image quality and quantitative accuracy of reconstructed optical properties are significantly improved with the proposed algorithm. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.24.5.051407](https://doi.org/10.1117/1.JBO.24.5.051407)]

Keywords: diffuse optical tomography; back-propagation neural network; image reconstruction, inverse problem.

Paper 180496SSRR received Aug. 13, 2018; accepted for publication Nov. 30, 2018; published online Dec. 19, 2018.

1 Introduction

Diffuse optical tomography (DOT) has shown a great potential for breast imaging^{1–9} and functional brain imaging,^{10–12} which use near-infrared light in the spectral range of 600 to 950 nm to quantify tissue optical (absorption and scattering) coefficients. There is a critical need to develop an efficient image reconstruction algorithm for DOT. Recovering the internal distribution of optical properties is a severely ill-posed and under-determined inverse problem, due to light propagation in highly scattering biological tissues and limited number of measurements,^{13,14} which makes image reconstruction challenging.

Although both linear and nonlinear reconstruction algorithms for DOT are available,¹⁴ considerable efforts have been made to develop various reconstruction algorithms to improve quantitative accuracy and image quality.^{14–22} To date, the ill-posedness of the inverse problem in DOT can be alleviated by employing a regularization technique, which utilizes a data fitting term together with a regularizer (L_2 or L_1 norm, etc.) to suppress the effect of measurement noise and modeling errors.²³

When the regularizer is a L_2 norm, the reconstruction algorithm becomes the well-known Tikhonov regularization method, which imposes restrictions on the L_2 norm of the optical properties²³ and is optimal when the distribution of optical properties subjects to Gaussian distribution.²⁴ The merits of Tikhonov regularization are simple, and easy to be implemented. However, the L_2 norm will oversmooth reconstructed

images and yield low-quality images by penalizing large values.²³ An alternative regularizer is total variation (TV) norm, which is the ideal choice when the distribution of optical properties is known to be piecewise constant.²⁴ Another possible regularizer is L_p norm ($0 < p \leq 1$), which poses a sparsity constraint on the optical properties. The quality of reconstructed images can be improved with the use of sparsity regularization.^{23,25}

The abovementioned algorithms are generally iterative reconstruction algorithms.¹⁴ These algorithms require heavy computation and large storage memory because the forward problem must be solved repeatedly, and an updated distribution of optical properties must be found at each iterative step.¹⁴ However, iterative reconstruction algorithms have limited capability in terms of reconstruction accuracy and image quality,²⁶ which are important for accurate diagnosis of breast cancers. In addition, how to accurately choose parameters, particularly regularization parameter, for iterative reconstruction algorithms needs to be further considered.²⁷

Recently, artificial neural networks with various network architectures, including deep convolutional neural network,^{28,29} generative adversarial networks,³⁰ and multilayer perceptron,³¹ have achieved significant improvements over existing iterative reconstruction methods in the quality of reconstructed images. It is likely that image recovery in DOT benefits from these important developments.

In this work, we investigated the feasibility and effectiveness of a back-propagation (BP) neural network (BPNN) to recover the distribution of optical properties in DOT. BPNN is a widely used neural network because it has many advantages. For

*Address all correspondence to Zhe Li, E-mail: lizhe1023@bjut.edu.cn; Kebin Jia, E-mail: kebinj@bjut.edu.cn

example, BPNN is simple, efficient at computing the gradient descent, and straightforward to implement. The basic procedure of BPNN includes the forward propagation of input data and the reverse transmission of output error.³² More detailed introduction about BPNN can be found in Refs. 33–39. The forward propagation of input data is to transmit input data from the input layer through a series of hidden layers toward the output layer, which builds the relationship between input data and output data. The reverse transmission of output error between the calculated and the ground true output is backward propagated from the output layer through the hidden layers to the first layer to adjust the connection weights and bias variables of neurons. By repeatedly applying this procedure, the output error is adjusted to an expected range. We validate the proposed method using simulation experiments and compare BPNN with the popular Tikhonov regularization. Our results demonstrate that our method provides higher accuracy and superior image quality than Tikhonov regularization in recovering a single inclusion or two closely spaced inclusions.

The remainder of the paper is organized as follows. Section 2 describes the light propagation model, BPNN, and evaluation metrics. Experimental results and comparisons are presented in Sec. 3. Finally, we present a discussion of results with our conclusions and future work in Sec. 4.

2 Methods

2.1 Forward Model

The light propagation in biological tissues can be modeled by the steady diffusion equation,^{13,14} which can be described as follows:

$$-\nabla \cdot D(r)\nabla\Phi(r) + \mu_a(r)\Phi(r) = q_0(r), \quad (r \in \Omega), \quad (1)$$

where Ω is the imaged object, $\Phi(r)$ is the photon fluence rate at position r , $D = 1/[3 * (\mu_a + \mu'_s)]$ is the diffusion coefficient (mm^{-1}), μ_a is the absorption coefficient (mm^{-1}), μ'_s is the reduced scattering coefficients (mm^{-1}), and $q_0(r)$ is the source term.

Here, the boundary condition used for Eq. (1) is Robin-type condition, which can be expressed as follows:^{13,14}

$$\Phi(r) + \frac{D}{\alpha} \hat{n} \cdot \nabla\Phi(r) = 0 \quad (r \in \partial\Omega), \quad (2)$$

where $\partial\Omega$ is the surface boundary of imaged object Ω , α is the boundary mismatch parameter, and \hat{n} is the outer normal on $\partial\Omega$.

When the distributions of μ_a and D are known, light measurements at the detectors can be calculated by solving Eqs. (1) and (2) based on the finite-element method,⁴⁰ which can be modeled with the following equation:

$$f(x) = \Phi^m, \quad (3)$$

where $x \in R^{2N}$ and $\Phi^m \in R^M$ represent the optical properties (μ_a and μ'_s) and the measurements at the detectors, respectively; N and M are the number of finite-element nodes and the number of boundary measurements, respectively; and $f(\cdot)$ is the forward operator that relates the unknown distribution of optical properties to the boundary measurements.

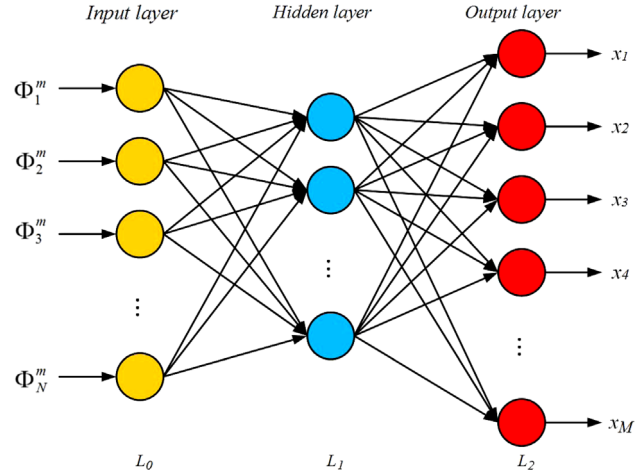


Fig. 1 The architecture of the three-layer BPNN used in DOT image reconstruction.

2.2 Back-Propagation Neural Network-Based Reconstruction

To improve the performances of iterative reconstruction algorithms in DOT, here we develop a reconstruction algorithm based on a BPNN. BPNN is divided into three types of layers: the input layer (L_0), the fully connected hidden layer (L_1), and the predictable output layer (L_2). The architecture of the three-layer BPNN used for DOT image reconstruction is shown in Fig. 1. We train the neural network from the boundary measurements to learn a DOT reconstruction. In this work, the reduced scattering coefficient (μ'_s) is assumed to be spatially constant and known and we recover only the absorption coefficient (μ_a). Therefore, for the network training, boundary measurements $\Phi_i (i = 1, \dots, M)$ (i.e., amplitude) are regarded as the input vector, which are generated by solving the forward model using open source software Nirfast,⁴⁰ and the ground true distribution of absorption coefficient $x_j (j = 1, \dots, N)$ is served as the expected output.

During the training process, the error between the predicted output and the true output is backward propagated from the output layer to the hidden layer to adjust the weights and biases in the opposite direction to the signal flow with respect to each individual weight.³² By repeatedly applying this procedure for each sample in the training set, the learning process can converge.

The BPNN-based reconstruction algorithm can be described as follows:³⁶

- Step 1: Randomly initialize the weights $w_{ij}^{[l]}$ and the bias variables $v_j^{[l]}$ at the l 'th ($l = 1, 2$) layer, and set the stop threshold δ and maximum number of iteration n_{\max} ;
- Step 2: Compute the output $o_j^{[l]}$ of the j 'th neuron at the l 'th ($l = 1, 2$) layer using Eq. (4):

$$o_j^{[l]} = g\left(\sum_{i=1}^{N^{[l]}} w_{ij}^{[l]} y_i^{[l-1]} + v_j^{[l]}\right), \quad (4)$$

where $g(\cdot)$ is the neuron activation function, $y_i^{[l]}$ is the input of the l 'th layer, $N^{[l]}$ is the number of the l 'th layer neurons. Here, $y^{[0]}$ is the input of the network, i.e., Φ^m .

Step 3: Calculate the mean square error for the output layer ($l = 2$) between the ground truth output d_j and the predicted output o_j according to Eq. (5):

$$\text{err}_j^{[2]} = -g'(o_j^{[2]})(d_j - o_j^{[2]}) \quad (j = 1, \dots, N^{[2]}) \quad (5)$$

and compute the mean square error in the hidden layer ($l = 1$) using Eq. (6):

$$\text{err}_j^{[1]} = g'(o_j^{[1]}) \sum_{k=1}^{N^{[2]}} \text{err}_k^{[2]} w_{jk}^{[2]}, \quad (6)$$

where $g'(\cdot)$ is the derivative of neuron activation function.

Step 4: Adjust the connection weights and biases between layers at the n 'th iteration based on Eqs. (7) and (8):

$$v_{ij}^{[l]}(n+1) = v_{ij}^{[l]}(n) + \eta \cdot \text{err}_j^{[l]} \quad (7)$$

$$w_{ij}^{[l]}(n+1) = w_{ij}^{[l]}(n) + \eta \cdot \text{err}_j^{[l]} \cdot o_i^{[l-1]} \quad (8)$$

where η is the learning rate, which controls the speed of adjusting neural network weights based on the gradient descent method.

Step 5: Go back to Step 2 if the mean squared error of the neural network output is larger than the given stop threshold δ , or n is smaller than n_{\max} , or the loss function is less than ε ; otherwise, the training processing will be terminated and output the weights and biases.

Step 6: Directly reconstruct the distribution of absorption coefficient by evaluating the trained network.

In our algorithm, the Tansig function is adopted as the activation function. Its formula is given in Eq. (9):

$$g(x) = \text{tansig}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad (9)$$

and its derivative is shown in Eq. (10):

$$g'(x) = 1 - [g(x)]^2. \quad (10)$$

2.3 Evaluation Metrics

The performance of the proposed algorithm is accessed with four evaluation metrics, including the absolute bias error (ABE), mean square error (MSE),^{41,42} peak signal-to-noise ratio (PSNR),⁴³ and structural similarity index (SSIM).⁴⁴ These parameters are defined as follows:

$$\text{ABE} = \frac{\sum_{i=1}^N |x_{\text{true}(i)} - x_{\text{recon}(i)}|}{N}, \quad (11)$$

$$\text{Var} = \frac{\sum_{i=1}^N |x_{\text{true}(i)} - x_{\text{recon}}|}{N}, \quad (12)$$

$$\text{MSE} = \text{ABE}^2 + \text{Var}, \quad (13)$$

$$\text{PSNR} = 10 \log_{10} \left\{ \frac{[\max(x_{\text{recon}})]^2}{\text{MSE}} \right\}, \quad (14)$$

$$\text{SSIM} = \frac{(2\bar{x}_{\text{true}}\bar{x}_{\text{recon}} + c_1)(2\sigma_{\text{true, recon}} + c_2)}{(\bar{x}_{\text{recon}}^2 + \bar{x}_{\text{recon}}^2 + c_1)(\sigma_{\text{true}}^2 + \sigma_{\text{recon}}^2 + c_2)}, \quad (15)$$

where $x_{\text{true}(i)}$ and $x_{\text{recon}(i)}$ are the true and reconstructed absorption coefficients at the finite node i , respectively; \bar{x}_p ($p = \text{true}$ or recon) and σ_p ($p = \text{true}$ or recon) are the mean and standard derivation for the ground true ($p = \text{true}$) or reconstructed ($p = \text{recon}$) absorption coefficients, respectively; $\sigma_{\text{true, recon}}$ is the covariance between the ground true and the reconstructed absorption coefficients, and c_1, c_2 are stabilization constants used to prevent division by a small denominator;⁴⁴ N is the above mentioned number of finite-element nodes. The ABE and MSE are used to compare the accuracy of reconstructed images. The PSNR (unit: dB) is used to compare the restoration of the images, without depending strongly on the image intensity scaling.¹⁵ SSIM is used to measure the similarity between the true and the reconstructed images, and an SSIM value of 1.0 refers to identical images. We expect lower ABE and MSE, while higher PSNR and SSIM, which show better performance.

3 Results

3.1 Data Preparation

A 2-D circular phantom with a diameter of 80 mm was used to generate dataset. It was discretized into 2001 finite-element nodes and 3867 triangular elements. The absorption coefficient (μ_a) and the reduced scattering coefficient (μ_s') of the phantom were 0.01 mm^{-1} and 1.0 mm^{-1} , respectively. A total of 16 sources and 16 detectors were uniformly arranged along the circumference of the phantom. For each source illumination, data were collected at the remaining 15 detector locations, thus leading to a total of 240 (16×15) measurements. To generate simulation datasets, the phantom includes circular inclusions associated with varied sizes, locations, and absorption coefficients. Initially, an inclusion with the diameter of 6, 8, or 10 mm was randomly placed at different locations over the phantom. In this case, the absorption coefficients of the inclusions were varied from 0.015 to 0.08 mm^{-1} , leading to 17075 geometries. Next, two inclusions, which have the same radius of 8 mm, were placed at different edge-to-edge distances over the phantom. In this case, the phantoms were assigned different absorption coefficients ($0.015, 0.02, 0.04, 0.06, \text{ or } 0.08 \text{ mm}^{-1}$) to each of its inclusions, leading to 5015 geometries. In all cases, the reduced scattering coefficients of inclusions were assumed to be 1 mm^{-1} , which were the same as those of the background. Examples of geometries of generating data are shown in Fig. 2. Software Nifrast was used to generate the simulation data,⁴⁰ and 2% random Gaussian noise was added to the measurement data. A total of 22,090 samples which are data pairs contained input data and desired output data were obtained, and were separated into training, validation, and testing datasets. About 20,000 samples were used for training, 1045 for validation, and 1045 for testing.

In the following experiments, the neural network has three layers: an input layer (240 neurons), a fully connected hidden layer, and an output layer (2001 neurons). To determine the number of neurons in the hidden layer, an empirical formula that has been introduced in Ref. 45 was used. The formula is given by

$$s = \sqrt{m(n+2)} + 1, \quad (16)$$

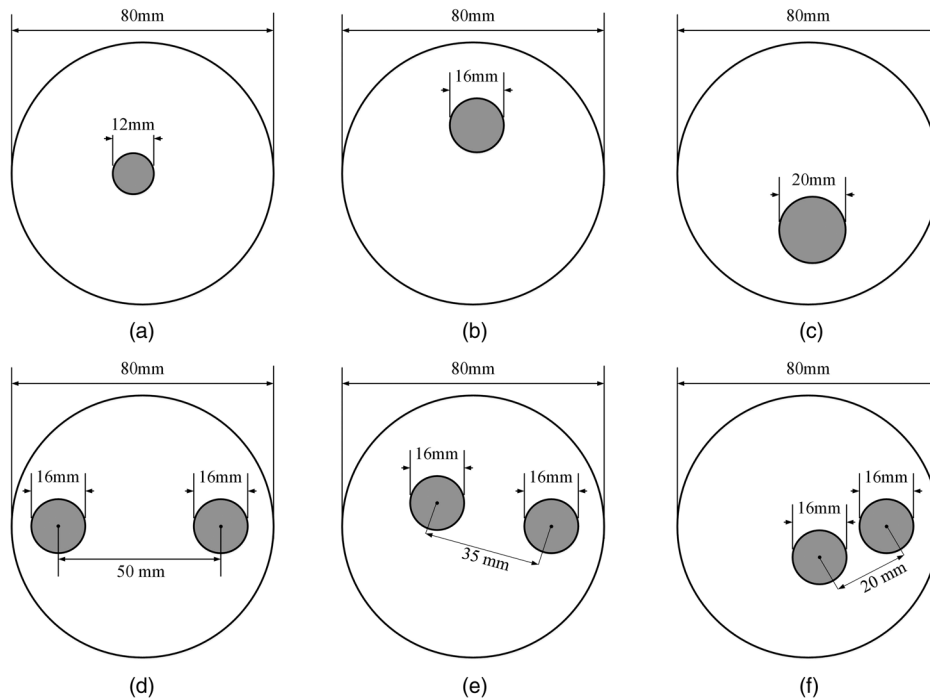


Fig. 2 Examples of geometries of generating datasets. (a)–(c) The examples of geometry with a single inclusion that has different locations and sizes. (d)–(f) The examples of geometries with two inclusions that have different edge-to-edge distance and locations. The absorption coefficients of inclusions in each geometry are varied from 0.015 to 0.08 mm^{-1} .

where s is the number of the hidden layer units, m and n are the numbers of input neurons and output neurons. In our experiments, the values of m and n are 240 and 2001, respectively. According to Eq. (16), the calculated number of neurons in hidden layer is 694.3. In our experiments, the number of neurons in the hidden layer was set to 695. The learning rate η , the stopping threshold δ , the maximum number of iteration and the threshold ε were set to be 10, 1×10^{-5} , 50,000 and 2×10^{-5} , respectively. BPNN was trained for 16,000 epochs to minimize the MSE between the true and the recovered images. It took about 26 h to train the BP neural network. The experiments ran on a personal computer with Intel Core i7 CPU at 2.8 GHz and 8 GB RAM.

For the purpose of comparison, we also performed Tikhonov regularization-based DOT reconstruction based on the Nirfast software.⁴⁰ Tikhonov regularization-based DOT reconstruction is achieved using a least squares (LS) minimization technique, which is solved in the Levenberg–Marquardt procedure.⁴⁰ The objective function in the Tikhonov regularization-based reconstruction algorithm typically consists of a data fidelity term of weighted LS and a regularization term of L_2 norm, balanced by a regularization parameter λ . For λ at the k 'th iteration, it was setting as $\lambda_k = 10 * \max[\text{diag}(J_k^T J_k)]$, where J_k is the Jacobian matrix at the k 'th iteration. The stopping criterion for Tikhonov regularization is defined such that the algorithm stops when the change in the difference between the forward data and the reconstructed data of two successive iterations is less than 2% or the maximum number of iteration (50) is reached. The initial regularization parameter is set to be 10. More detailed information about Tikhonov regularization-based DOT reconstruction can be found in Ref. 40.

3.2 Experimental Results

In this subsection, we provide numerical simulations to illustrate recovered results using BPNN and compare it with widely used Tikhonov regularization-based reconstruction method. Figure 3 shows some examples of recovered absorption coefficient using the two algorithms in the case of single inclusion. In Fig. 3, the size, the location, and the absorption coefficient of inclusions are varied. The ground true images are shown in the top row of Fig. 3, the reconstructed images using Tikhonov regularization and BPNN are given in the second and the third rows of Fig. 3, respectively. The corresponding cross-section profiles through their centers of the inclusions and along the x axis are plotted in the last row of Fig. 3. From the last row of Fig. 3, we can see that the maximum values of recovered μ_a using Tikhonov regularization are much higher compared to their ground truths, the recovered μ_a using BPNN matched with their true values. The quantitative comparisons for the five cases in Fig. 3 are listed in Table 1. Compared with Tikhonov regularization, the values of ABE and MSE obtained using BPNN are significantly reduced, and the values of PSNR and SSIM are greatly improved. As an example, the values of ABE and MSE in Fig. 3(a) are reduced by 80% and 61%, respectively, which show that BPNN can yield higher reconstruction accuracy. It is evident that BPNN provides high-quality images with less artifacts in the background than those of Tikhonov regularization. Therefore, BPNN can have a PSNR gain of about 4.4 dB over Tikhonov regularization while the higher value of SSIM is obtained. The value of SSIM is improved by 424%, which indicates that the recovered image is nearly the same with the ground truth image. The similar results can also be observed in other images of Fig. 3. These results show that BPNN outperforms Tikhonov regularization in terms of higher accuracy and better

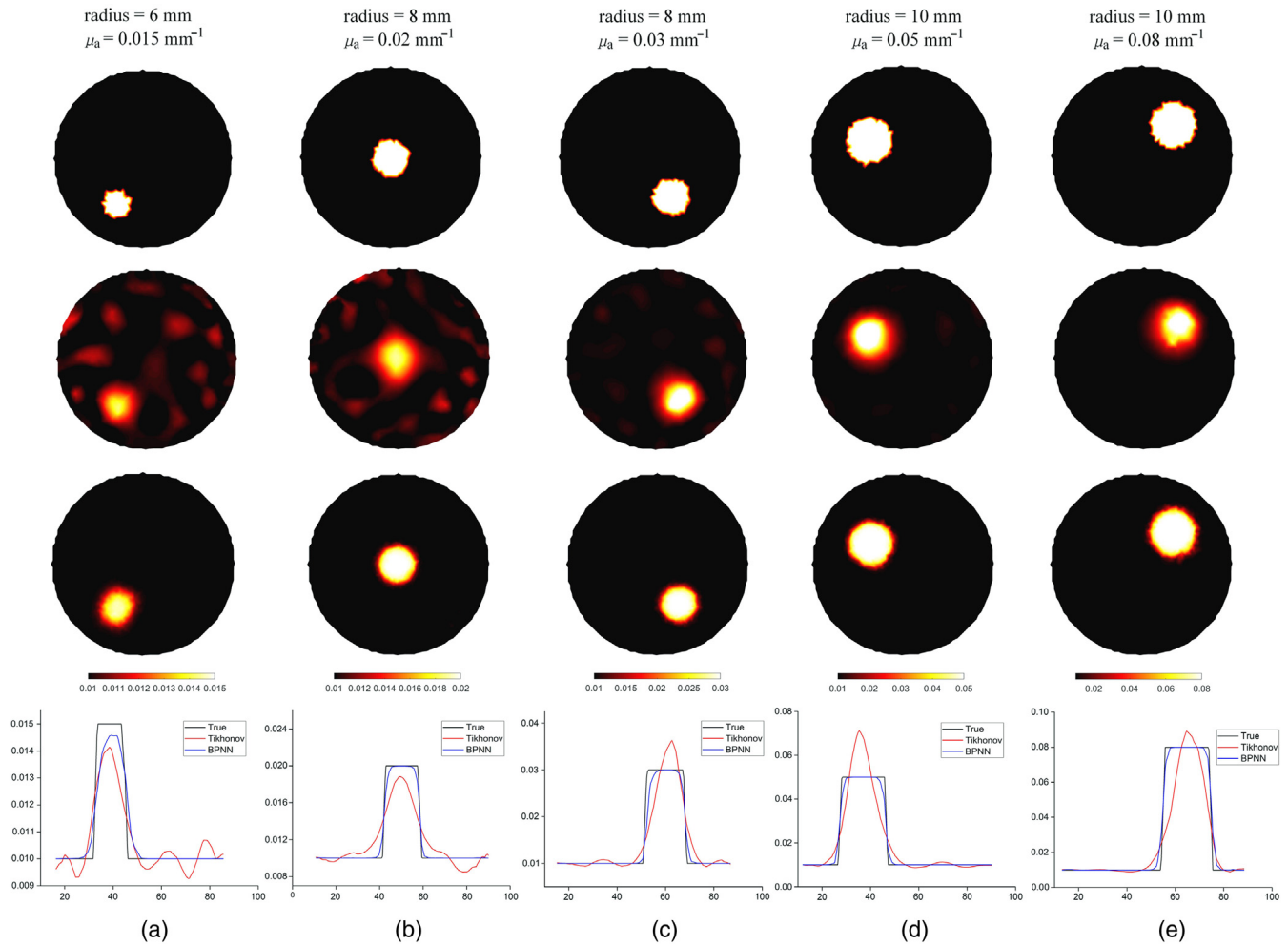


Fig. 3 (a)–(e) Reconstructed images of a single inclusion with different sizes and different values of absorption coefficients. The first row is the true images, the second and third rows are the recovered images using Tikhonov regularization and BPNN, respectively. The last row is the corresponding profiles through the center of inclusions and along x axis. The sizes and the true values of absorption coefficients for each case are shown at the top of the figure. The reconstructed images of each column are shown at the same scale.

Table 1 The quantitative comparisons presented in Fig. 3.

Metric	Method	Fig. 3(a)	Fig. 3(b)	Fig. 3(c)	Fig. 3(d)	Fig. 3(e)
ABE	Tikhonov	4.35×10^{-4}	8.52×10^{-4}	8.70×10^{-4}	1.69×10^{-3}	2.50×10^{-3}
	BPNN	8.51×10^{-5}	1.00×10^{-4}	1.77×10^{-4}	4.38×10^{-4}	8.60×10^{-4}
MSE	Tikhonov	3.75×10^{-7}	1.62×10^{-6}	3.35×10^{-6}	2.06×10^{-5}	6.19×10^{-5}
	BPNN	1.45×10^{-7}	2.30×10^{-7}	9.78×10^{-7}	5.63×10^{-6}	1.90×10^{-5}
PSNR (dB)	Tikhonov	27.37	23.43	26.10	24.44	21.58
	BPNN	31.75	32.40	29.64	26.47	25.28
SSIM	Tikhonov	0.17	0.20	0.45	0.75	0.84
	BPNN	0.89	0.97	0.96	0.96	0.96

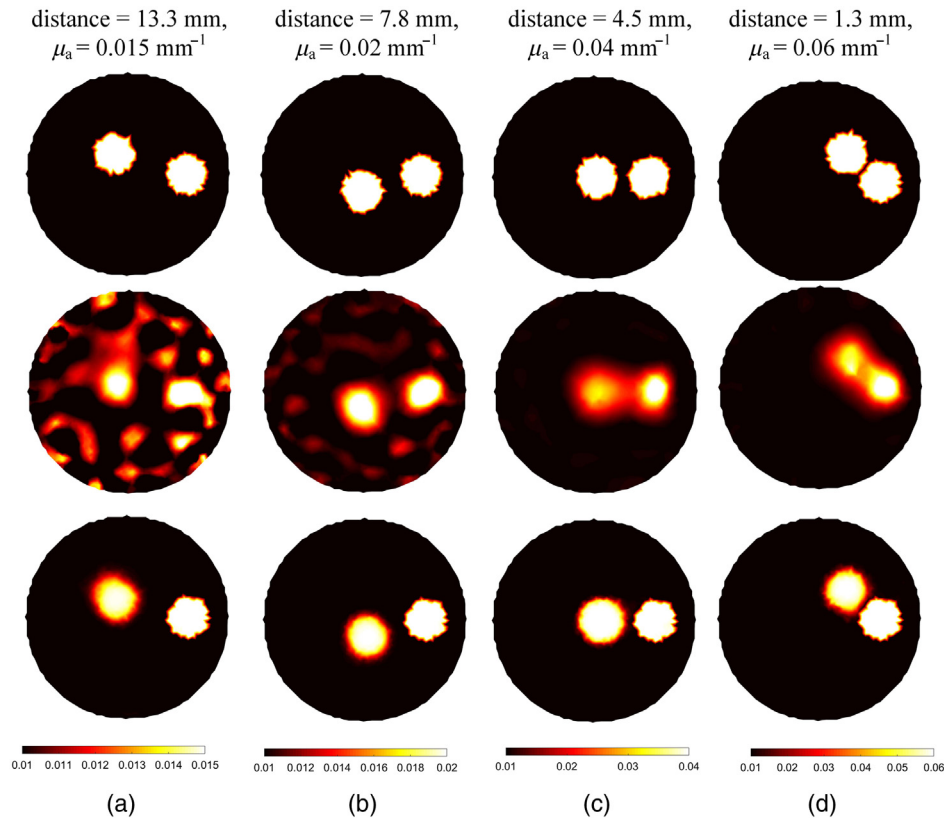


Fig. 4 (a)–(d) Examples of reconstructed images for increased intensity of absorption coefficient but with decreased edge-to-edge distance in the case of two inclusions. The first row is the ground truth images. From the second row to the last row, the images are reconstructed using Tikhonov regularization and BPNN, respectively. The edge-to-edge distance and the true values of absorption coefficients for each case are shown at the top of the figure. The reconstructed images of each column are shown at the same scale.

image quality. In addition, the results also show that the image quality obtained by Tikhonov regularization is improved with increment of size and absorption coefficient of inclusions. By contrast, BPNN can always obtain robust reconstruction results.

Figure 4 shows the capability of BPNN to recover images with two inclusions. And the corresponding quantitative results are compiled in Table 2. The two inclusions are observable and

reconstructed with their centers at the correct positions for the two algorithms. But it can be seen that the shapes and the edges of the inclusions can be clearly observed by BPNN even when the edge-to-edge distance is 1.3 mm. However, the absorption coefficients of the inclusions far away from the boundary were underestimated and the images were somehow distorted for the Tikhonov regularization-based reconstruction method. For each

Table 2 The quantitative comparisons in the case of two inclusions presented in Fig. 4.

Metric	Method	Fig. 4(a)	Fig. 4(b)	Fig. 4(c)	Fig. 4(d)
ABE	Tikhonov	1.30×10^{-3}	1.07×10^{-3}	2.07×10^{-3}	2.84×10^{-3}
	BPNN	1.28×10^{-4}	2.10×10^{-4}	5.50×10^{-4}	6.04×10^{-4}
MSE	Tikhonov	2.83×10^{-6}	2.74×10^{-6}	1.89×10^{-5}	4.90×10^{-5}
	BPNN	3.13×10^{-7}	1.16×10^{-6}	1.01×10^{-5}	9.29×10^{-6}
PSNR (dB)	Tikhonov	22.55	25.35	21.76	21.92
	BPNN	28.55	25.39	22.0	25.88
SSIM	Tikhonov	0.09	0.36	0.67	0.78
	BPNN	0.92	0.92	0.91	0.97

evaluation metric, we draw a bar plot of test images in Fig. 5. As the higher accuracy of BPNN can be clearly observed in Fig. 5 and Table 2, we do not compare the cross-sections of different images here. Similar to the previous simulation results of one inclusion, BPNN reduces the artifacts and has offered more than 55.0%, 46.6%, and 16.7% improvement in ABE, MSE, and SSIM, respectively, compared with Tikhonov regularization-based reconstruction algorithm. The PSNRs in Figs. 4(b) and 4(c) are slightly higher than those of the Tikhonov regularization-based reconstruction algorithm because the absorption coefficients reconstructed by Tikhonov regularization have been overestimated with peak values of 0.03 and 0.05 mm⁻¹, respectively. By contrast, the peak values obtained by BPNN are 0.02 and 0.04 mm⁻¹, which are the same as the true targeted values. For Fig. 4, BPNN can have an average PSNR gain of about 2.56 dB over Tikhonov regularization. Overall, BPNN can obtain better performance.

Here, we reported the results of 1045 samples, which was randomly selected from the dataset to further evaluate the performance of our algorithm. We use the mean and standard deviation (SD) of ABE, MSE, PSNR, and SSIM to evaluate the performances of the two algorithms. The boxplots for the statistical results are presented in Fig. 6. And the corresponding quantitative comparisons are shown in Table 3.

For BPNN, the ABE had a mean value (SD) of 3.41×10^{-4} (2.50×10^{-4}); for Tikhonov regularization, the ABE had a mean value (SD) of 1.50×10^{-3} (9.26×10^{-4}). The accuracy improvement is 77.3% compared with Tikhonov regularization. The MSE of BPNN has an average of 5.97×10^{-6} , while the average for Tikhonov regularization is 2.30×10^{-5} . The improvement in the average MSE of BPNN with respect to Tikhonov regularization is 74.0%. Furthermore, the average PSNR of BPNN is improved from 24.34 to 27.79 dB compared with Tikhonov regularization. The average value of SSIM (0.91) obtained by BPNN is significant as compared to the average value of SSIM (0.46) produced by Tikhonov regularization. For every one of these metrics, BPNN performs better than Tikhonov regularization being compared.

A student's *t*-test is used to determine whether there is statistical significance of the improvement between the evaluation metrics of Tikhonov regularization and BPNN. Significance is achieved at the 95% confidence interval using a two-tailed distribution. The corresponding *p*-values are also listed in Table 3. Statistically significant differences (*p* values <0.001) were found in the four evaluation metrics for BPNN versus Tikhonov regularization results, which further confirms that BPNN achieves more stable and effective performance than Tikhonov regularization.

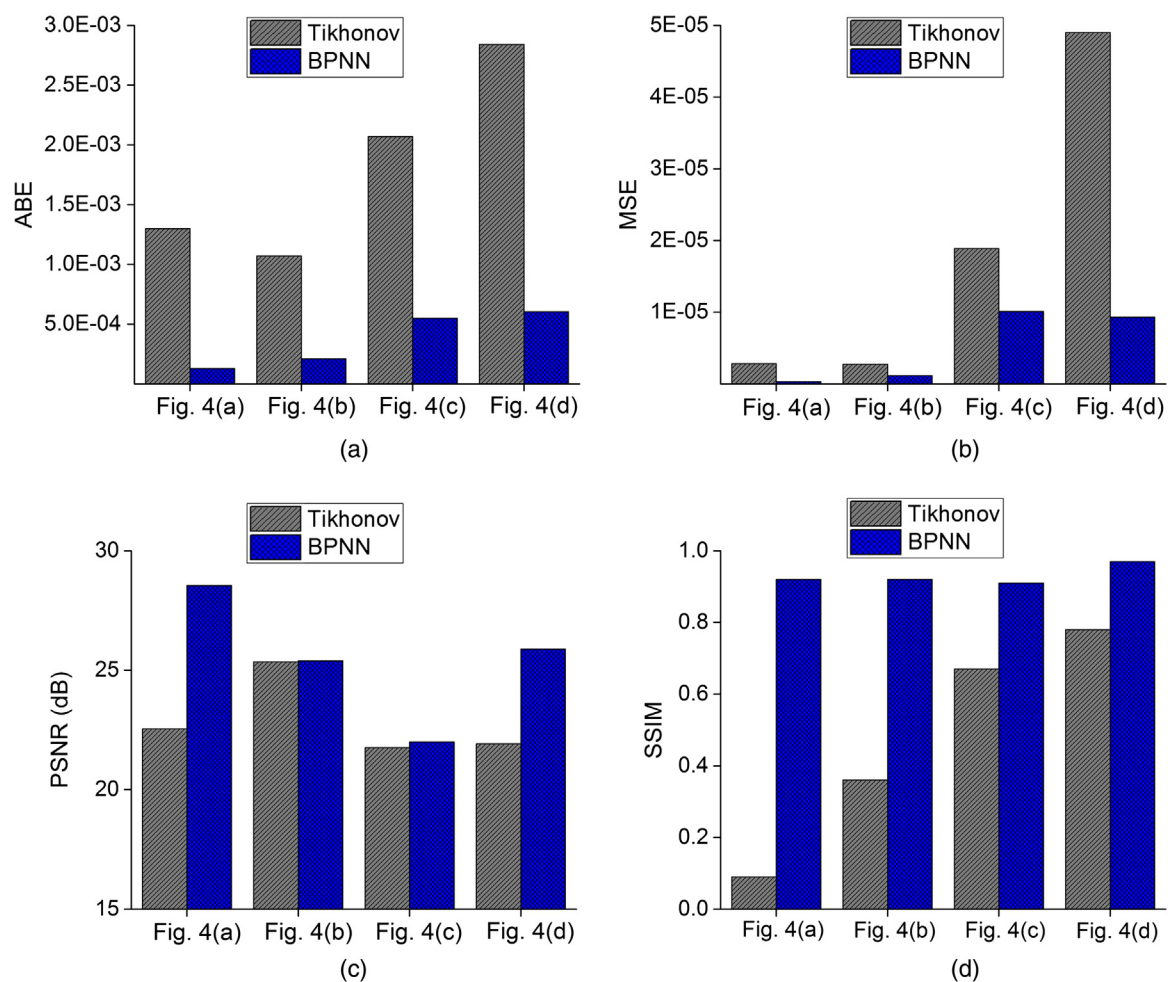


Fig. 5 Comparisons of evaluation metrics for reconstructed images of two inclusions using Tikhonov regularization and BPNN. (a) ABE; (b) MSE; (c) PSNR; and (d) SSIM.

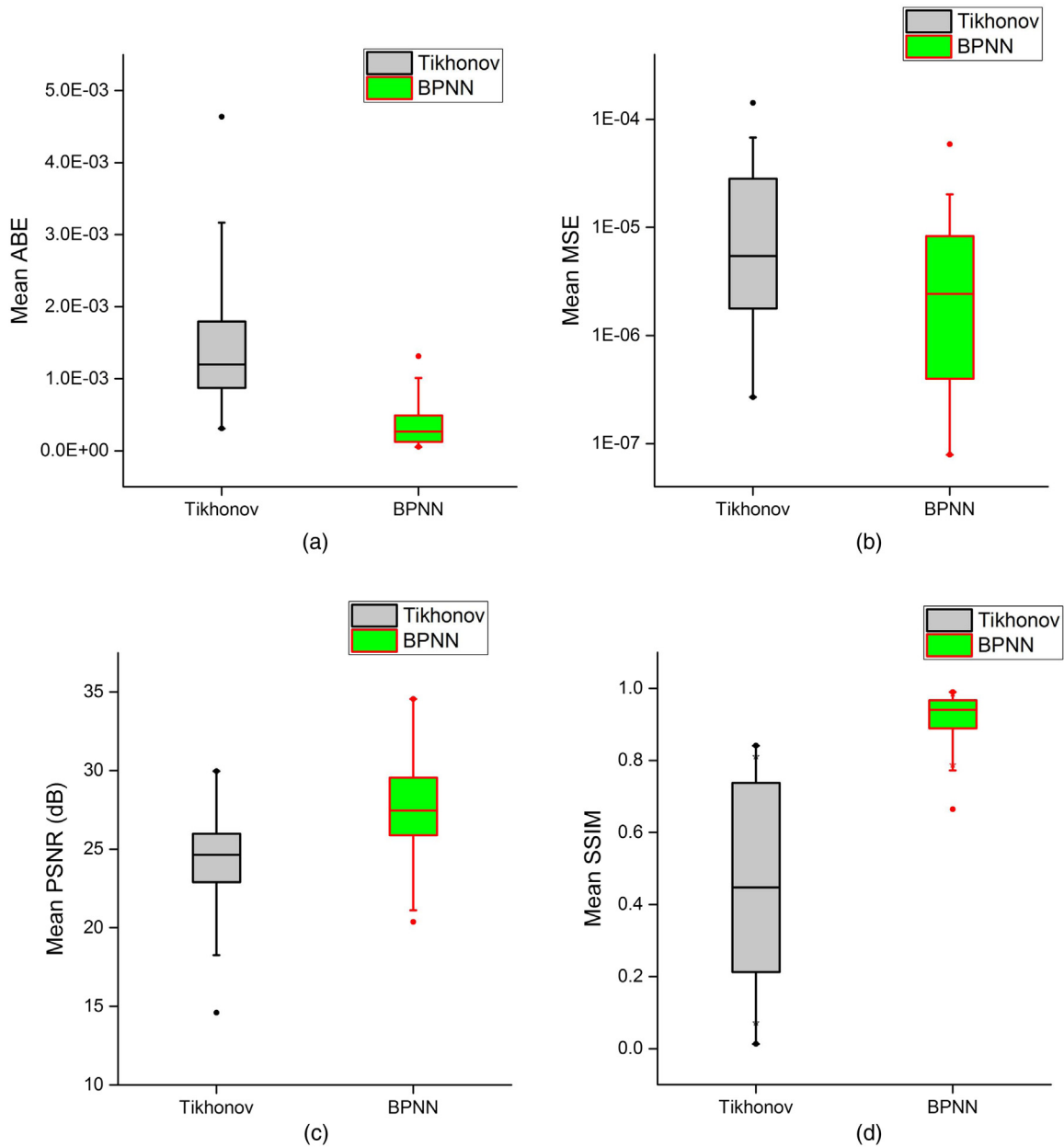


Fig. 6 The boxplots for the statistical results ($N = 1045$). (a) ABE; (b) MSE; (c) PSNR; and (d) SSIM.

Table 3 Mean \pm SD of ABE/MSE/PSNR/SSIM for $N = 1045$.

	ABE	MSE	PSNR	SSIM
Tikhonov	$1.50 \times 10^{-3} \pm 9.26 \times 10^{-4}$	$2.30 \times 10^{-5} \pm 3.29 \times 10^{-5}$	24.34 ± 2.43	0.46 ± 0.26
BPNN	$3.41 \times 10^{-4} \pm 2.50 \times 10^{-4}$	$5.97 \times 10^{-6} \pm 8.07 \times 10^{-6}$	27.79 ± 2.70	0.91 ± 0.06
<i>P</i> -value	<0.001*	<0.001*	<0.001*	<0.001*

*Significant values are marked.

4 Discussion and Conclusion

Iterative reconstruction algorithms with regularization have become the dominant approach for solving DOT inverse problem over the past few decades. However, it remains difficult to provide high-quality images. In this study, we explored

using a BPNN to recover optical properties to improve the reconstruction accuracy and image quality of DOT. This was motivated by the fact that popular Tikhonov regularization-based reconstruction algorithms tend to produce oversmoothed images, which leads to poor reconstruction accuracy and bad

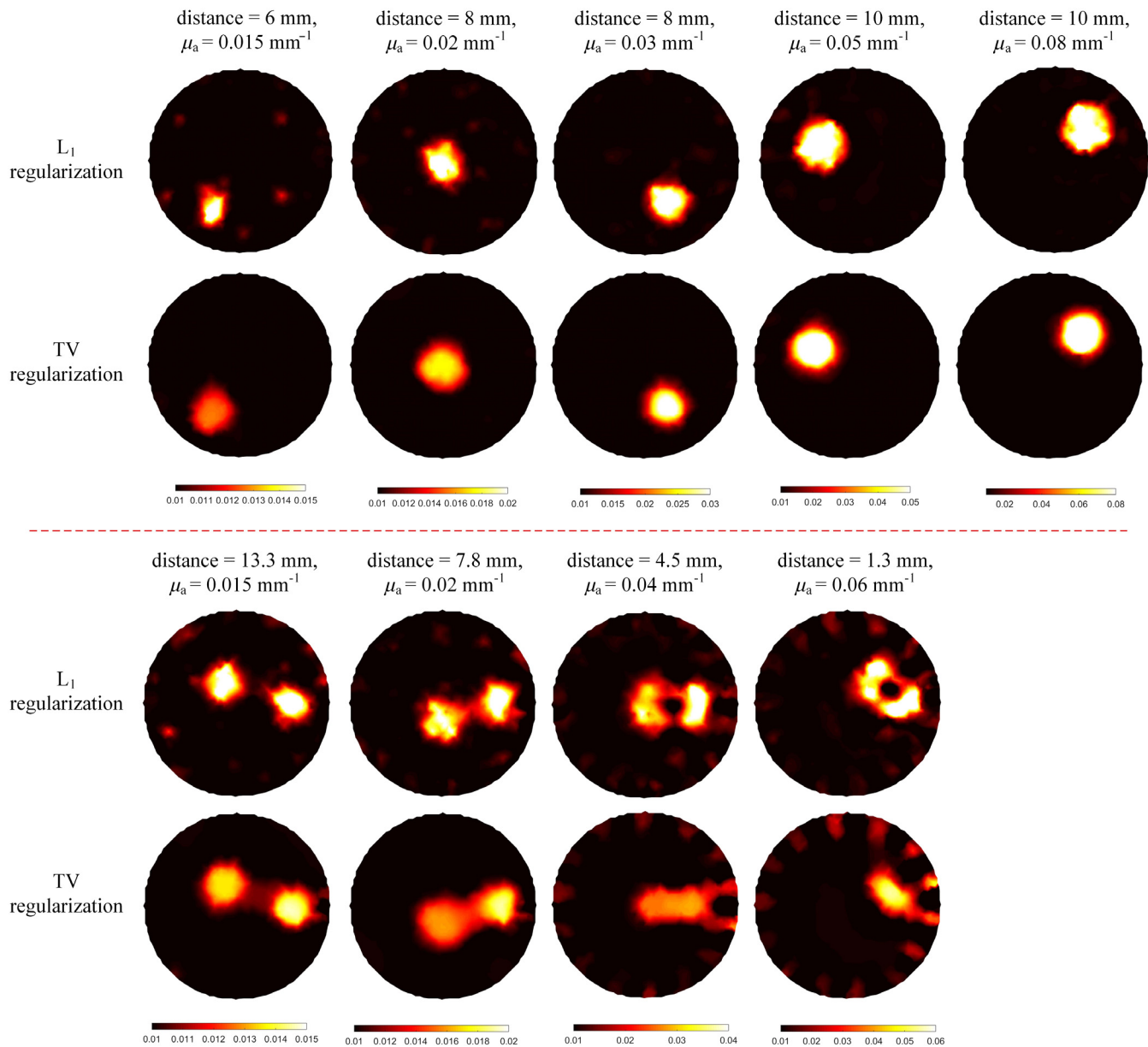


Fig. 7 Reconstructed results by both L_1 and TV regularization algorithms, along with the corresponding reconstructions in Figs. 3 and 4. Top images: reconstructed images with a single inclusion. Bottom images: reconstructed images with two inclusions.

image quality. The superior performance of the proposed algorithm was presented with numerical simulation experiments. Our results indicate that significant improvements of accuracy and image quality can be achieved by the proposed algorithm when compared with the Tikhonov regularization-based algorithm. Qualitative analysis demonstrated that our method can outperform Tikhonov regularization up to 77.3%, 74.0%, 14.2%, and 97.8% in terms of ABE, MSE, PSNR, and SSIM, respectively.

Furthermore, we compared the reconstructed results of the proposed method to those of the L_1 and TV regularized reconstruction algorithms. The examples of reconstructed images are shown in Fig. 7. For the L_1 regularized reconstruction algorithm, it was solved with the GPSR algorithm.⁴⁶ For the TV regularized reconstruction algorithm, it was solved by the Split Bregman algorithm.⁴⁷ As for the regularization

parameters used in the two reconstruction algorithms, their values were the same and were set to 0.01, which was tuned manually to get the best performance. In the case of single inclusion, the L_1 - or TV-based reconstruction algorithm can obtain better images than the Tikhonov regularization-based reconstruction algorithm in terms of less artifacts. In the case of two inclusions, decreasing the edge-to-edge distance between inclusions leads to degraded image quality for the L_1 - and TV-based reconstruction algorithms. When the edge-to-edge distance of inclusions decreased to 4.5 mm, the two inclusions could not be discriminated for the TV-based reconstruction algorithm. Compared to regularization-based reconstruction algorithms (L_2 , L_1 , and TV), our algorithm performs the best in reconstructing DOT images.

It is possible to obtain high-quality DOT images by training a neural network, even when iterative reconstruction algorithms

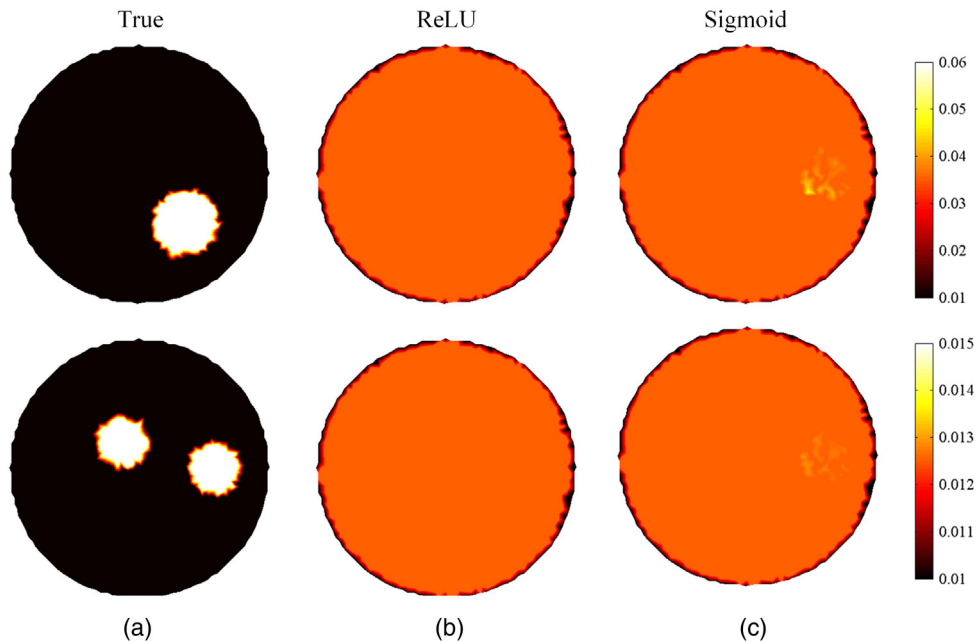


Fig. 8 The reconstructed images using different activation functions. (a) True images; (b) and (c) reconstructed images with activation functions of ReLU and sigmoid, respectively.

underperform. Compared to iterative reconstruction algorithms, BPNN-based reconstruction algorithm is capable of: (1) improving the reconstruction accuracy with relatively stable performance; (2) enhancing the image quality with fewer image artifacts in the background; and (3) improving the spatial resolution.

Currently, computational speed is still an open active research area in DOT. To the best of our knowledge, the widely used iterative reconstruction algorithms in DOT require to solve the forward model and to calculate the Jacobian matrix at each iteration.⁴⁰ Therefore, the computational speed is relatively slow. For example, the computational time for each reconstruction in this study is about 1 to 2 min for Tikhonov regularization. For this reason, an algorithm that can fast reconstruct the distribution of optical properties is preferred. Although it takes a rather longer time to train the neuron network for the proposed algorithm, the training is implemented off-line. Once the training is finished, the time for the reconstruction is in a few seconds, which is practically useful for *in vivo* data.

The effects of different activation functions, including ReLU, Sigmoid, and Tansig, were also examined. The representative results are shown in Fig. 8. Figure 8 shows that the distribution of absorption properties cannot be accurately reconstructed when using either ReLU or sigmoid as the activation function. The mean values of SSIM were 0.02 and 0.03 when using ReLU and sigmoid as the activation function, respectively. There are significant differences between the reconstructed and true images. By contrast, better images can be achieved with the activation function of Tansig and the mean SSIM is 0.91. However, it is still unclear why the activation function of Tansig works for DOT reconstruction.

How to determine the number of neurons in the hidden layer needs to be investigated. Except for Eq. (16), one of the typical equations which can be used to determine the number of neurons in the hidden layer is $\log_2(n)$,⁴⁸ where n is the number of input neurons. In our experiments, n is 240. Therefore, $\log_2(n)$ is about 8. We trained our network with eight hidden neurons.

Using the trained network, DOT image reconstruction was performed. The examples of reconstructed images are shown in Fig. 9. Figure 9 shows that the distribution of absorption coefficient cannot be accurately reconstructed. The reason is that using too few neurons in the hidden layer will result in underfitting.

The effects of the learning rate on the performance of BPNN were also analyzed. The learning rate was set to 0.01, 0.1, 1, 10, or 50. The optimal learning rate was the one with the lowest validation loss. In the experiments, the activation function was fixed as Tansig. When the learning rate was set to 50, the validation loss fluctuated over epochs. Therefore, the learning rate of 50 was discarded. The validation losses were 0.15, 0.09, 0.04, and 0.02, corresponding to the learning rates, 0.01, 0.1, 1, and 10, respectively. When learning rate was 10, the lowest validation loss was obtained. We also note that a higher learning rate effectively speeds up the convergence for the training procedure. Therefore, the learning rate was set to 10 in our experiments.

Note that several recent publications have applied various deep learning architectures for solving inverse problems.^{49–51} For example, a recent work by Sun et al. has shown that a convolutional neural network can be applied to perform image reconstruction under multiple scattering problem in diffraction tomography.⁴⁹ However, our focus is on diffuse optical tomography. Recently, an artificial neural network-based approach is developed to estimate the inclusion location, then the estimated inclusion location is used as a-priori knowledge in DOT reconstruction.⁵⁰ Therefore, the approach is not to reconstruct DOT images directly by learning an artificial neural network. Deep learning technique has been recently applied to reconstruct DOT images, and its superiority has been indicated by comparing with an analytic technique.⁵¹ The average value of SSIM they reported is 0.46, which is relatively low. By contrast, a higher value of SSIM is obtained by our algorithm, and its value is 0.91. A future study would be to compare the performances of the two algorithms in the same datasets.

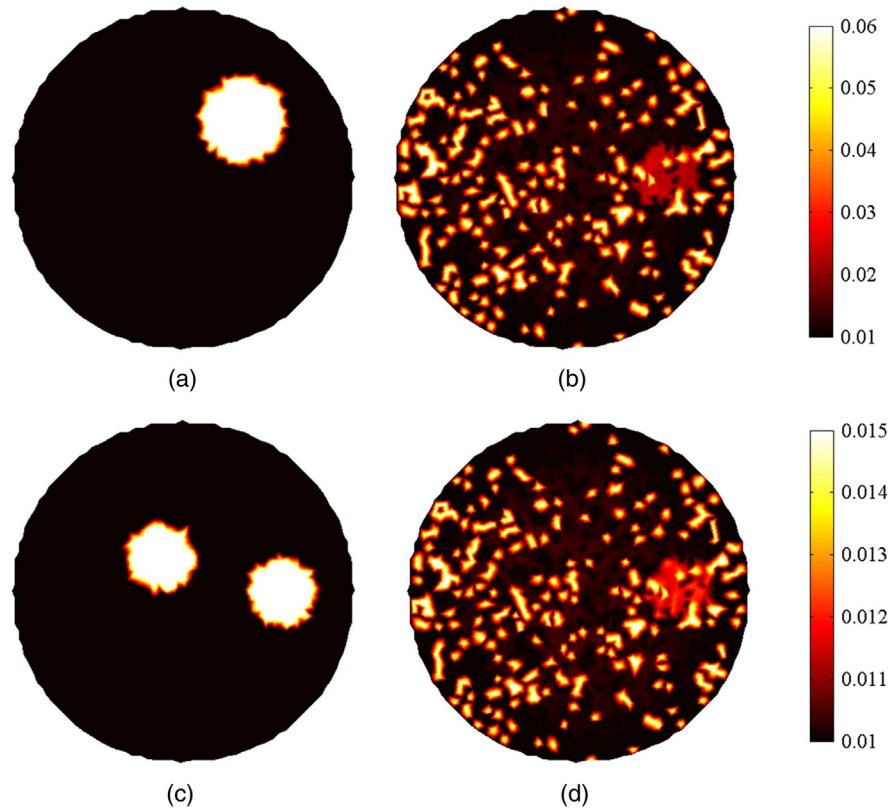


Fig. 9 The examples of reconstructed images using the neural network with eight hidden neurons. The first and second rows are the results of single inclusion and two inclusions, respectively. (a) and (c) The true images; (b) and (d) the reconstructed images.

Although the proposed BPNN method has achieved promising results, our work could still be considered as a preliminary attempt of applying neural network in DOT; its application in DOT is still very challenging. The performance of BPNN depends on the training data; however, it would be difficult to acquire a sufficient number of real data for training in patient studies. An available strategy is to create training data pairs from breast geometries with known optical properties and the real data is for evaluation. Breast geometries can be obtained from breast MRI images. Since breast geometries have different sizes and shapes, the finite-element meshes generated from breast geometries are different, which leads to different BPNN architectures. To deal with the problem, pixel basis provides a solution.⁴⁰ Each finite-element mesh is mapped to the same pixel basis for network training. Therefore, the trained BPNN will be a universal network. Certain challenges remain which are the subject of further study, including the effect of heterogeneity in the background region as well as performance evaluation using clinical patient data.

Disclosures

The authors have no relevant financial interests in the manuscript.

Acknowledgments

This work has been sponsored by the National Science Foundation of China under Grant No. 81871394, the Beijing Municipal Education Committee Science Foundation under Grant No. KM201810005030, and Beijing Laboratory of Advanced Information Networks under Grant No. 040000546618017.

References

1. J. Feng et al., "Addition of T2-guided optical tomography improves noncontrast breast magnetic resonance imaging diagnosis," *Breast Cancer Res.* **19**, 117 (2017).
2. Q. Zhu et al., "Assessment of functional differences in malignant and benign breast lesions and improvement of diagnostic accuracy by using US-guided diffuse optical tomography in conjunction with conventional US," *Radiology* **280**(2), 387–397 (2016).
3. C. Xu et al., "Ultrasound-guided diffuse optical tomography for predicting and monitoring neoadjuvant chemotherapy of breast cancers: recent progress," *Ultrason. Imaging* **38**(1), 5–18, (2016).
4. V. Ntziachristos et al., "Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement," *Proc. Natl. Acad. Sci. U. S. A.* **97**(6), 2767–2772 (2000).
5. B. Brooksby et al., "Imaging breast adipose and fibroglandular tissue molecular signatures by using hybrid MRI-guided near-infrared spectral tomography," *Proc. Natl. Acad. Sci. U. S. A.* **103**(23), 8828–8833 (2005).
6. M. A. Mastanduno et al., "Sensitivity of MRI-guided near-infrared spectroscopy clinical breast exam data and its impact on diagnostic performance," *Biomed. Opt. Express* **5**, 3103–3115 (2014).
7. S. Jiang et al., "Predicting breast tumor response to neoadjuvant chemotherapy with diffuse optical spectroscopic tomography prior to treatment," *Clin. Cancer Res.* **20**(23), 6006–6015 (2014).
8. M. Mastanduno et al., "MR-guided near-infrared spectral tomography increases diagnostic performance of breast MRI," *Clin. Cancer Res.* **21**(17), 3906–3912 (2015).
9. V. Ntziachristos et al., "MRI-guided diffuse optical spectroscopy of malignant and benign breast lesions," *Neoplasia* **4**, 347–354 (2002).
10. C. Chen et al., "Diffuse optical tomography enhanced by clustered sparsity for functional brain imaging," *IEEE Trans. Med. Imaging* **33**(12), 2323–2331 (2014).
11. M. A. Franceschini and D. A. Boas, "Noninvasive measurement of neuronal activity with near-infrared optical imaging," *Neuroimage* **21**(1), 372–386 (2004).

12. C. W. Lee, R. J. Cooper, and T. Austin, "Diffuse optical tomography to investigate the newborn brain," *Pediatr. Res.* **82**, 376–386 (2017).
13. S. R. Arridge, "Optical tomography in medical imaging," *Inverse Prob.* **15**(2), R41–R93 (1999).
14. S. R. Arridge and J. C. Schotland, "Optical tomography: forward and inverse problems," *Inverse Probl.* **25**(12), 123010 (2009).
15. J. Feng et al., "Impact of weighting function on reconstructed breast images using direct regularization imaging method for near-infrared spectral tomography," *Biomed. Opt. Express* **9**(7), 3266–3283 (2018).
16. R. Baikejiang, W. Zhang, and C. Li, "Diffuse optical tomography for breast cancer imaging guided by computed tomography: a feasibility study," *J. X-Ray Sci. Technol.* **25**, 341–355 (2017).
17. J. Feng et al., "Multiobjective guided priors improve the accuracy of near-infrared spectral tomography for breast imaging," *J. Biomed. Opt.* **21**(9), 090506 (2016).
18. L. Zhang et al., "Direct regularization from co-registered anatomical images for MRI-guided near-infrared spectral tomographic image reconstruction," *Biomed. Opt. Express* **6**, 3618–3630 (2015).
19. Q. Fang et al., "Combined optical and X-ray tomosynthesis breast imaging," *Radiology* **258**(1), 89–97 (2011).
20. P. K. Yalavarthy et al., "Structural information within regularization matrices improves near infrared diffuse optical tomography," *Opt. Express* **15**, 8043–8058 (2007).
21. Z. Yuan and H. Jiang, "Image reconstruction scheme that combines modified Newton method and efficient initial guess estimation for optical tomography of finger joints," *Appl. Opt.* **46**(14), 2757–2768 (2007).
22. P. K. Yalavarthy et al., "Weight-matrix structured regularization provides optimal generalized least-squares estimate in diffuse optical tomography," *Med. Phys.* **34**(6), 2085–2098 (2007).
23. N. Cao, A. Nehorai, and M. Jacobs, "Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm," *Opt. Express* **15**(21), 13695–13708 (2007).
24. A. P. Gibson, J. C. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Phys. Med. Biol.* **50**, R1–R43 (2005).
25. J. Prakash et al., "Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction," *IEEE J. Sel. Top. Quantum Electron.* **20**(2), 74–82 (2014).
26. J. Prakash et al., "Model-resolution based basis pursuit deconvolution improves diffuse optical tomographic imaging," *IEEE Trans. Med. Imaging* **33**(4), 891–901 (2014).
27. Y. Zhao et al., "Optimization of image reconstruction for magnetic resonance imaging-guided near-infrared diffuse optical spectroscopy in breast," *J. Biomed. Opt.* **20**(5), 056009 (2015).
28. X. Yang et al., "Low-dose x-ray tomography through a deep convolutional neural network," *Sci. Rep.* **8**, 2575 (2018).
29. D. M. Pelt and K. J. Batenburg, "Fast tomographic reconstruction from limited data using artificial neural networks," *IEEE Trans. Image Process.* **22**(12), 5238–5251 (2013).
30. G. Yang et al., "DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Trans. Image Process.* **37**(6), 1310–1321 (2018).
31. B. Yang, L. Ying, and J. Tang, "Artificial neural network enhanced Bayesian PET image reconstruction," *IEEE Trans. Image Process.* **37**(6), 1297–1309 (2018).
32. C. Lv et al., "Levenberg-Marquardt backpropagation training of multi-layer neural networks for state estimation of a safety critical cyber-physical system," *IEEE Trans. Ind. Inf.* **14**(8) 3436–3446 (2018).
33. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature* **323**(6088), 533–536 (1986).
34. P. D. Wasserman, *Neural Computing: Theory and Practice*, Van Nostrand Reinhold, New York (1989).
35. R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Int. Joint Conf. on Neural Networks*, Vol. **1**, pp. 593–605 (1989).
36. S. Haykin, *Neural Networks*, Macmillan College Publishing Company, Inc., New York (1994).
37. I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *J. Microbiol. Methods* **43**(1), 3–31 (2000).
38. F. Amato et al., "Artificial neural networks in medical diagnosis," *J. Appl. Biomed.* **11**, 47–58 (2013).
39. E. Kuriscak et al., "Biological context of Hebb learning in artificial neural networks, a review," *Neurocomputing* **152**, 27–35 (2015).
40. H. Dehghani et al., "Near infrared optical tomography using NIRFAST: algorithm for numerical model and image reconstruction," *Commun. Num. Methods Eng.* **25**(6), 711–732 (2009).
41. B. W. Pogue et al., "Statistical analysis of nonlinearly reconstructed near-infrared tomographic images: part I-theory and simulations," *IEEE Trans. Med. Imaging* **21**(7), 755–763 (2002).
42. X. Song et al., "Statistical analysis of nonlinearly reconstructed near-infrared tomographic images: part II-experimental interpretation," *IEEE Trans. Med. Imaging* **21**(7), 764–772 (2002).
43. A. P. Cuadros and G. R. Arce, "Coded aperture optimization in compressive X-ray tomography: a gradient descent approach," *Opt. Express* **25**(20), 23833–23849 (2017).
44. Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
45. D. Gao and S. Wu, "An optimization method for the topological structures of feed-forward multi-layer neural networks," *Pattern Recognit.* **31**(9), 1337–1342 (1998).
46. M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Signal Process.* **1**(4), 586–597 (2007).
47. J. Feng et al., "Total variation regularization for bioluminescence tomography with the split Bregman method," *Appl. Opt.* **51**(19), 4501–4512 (2012).
48. R. P. Gorman and T. J. Sejnowski, "Analysis of hidden units in a layered network trained to classify sonar target," *Neural Networks* **1**, 75–89 (1988).
49. Y. Sun, Z. Xia, and U. S. Kamilov, "Efficient and accurate inversion of multiple scattering with deep learning," *Opt. Express* **26**(11), 14678–14688 (2018).
50. R. Patra and P. K. Dutta, "Improved dot reconstruction by estimating the inclusion location using artificial neural network," *Proc. SPIE* **8668**, 86684C (2013).
51. H. B. Yedder et al., "Deep learning based image reconstruction for diffuse optical tomography," in *Machine Learning for Medical Image Reconstruction*, F. Knoll, A. Maier, and D. Rueckert, Eds., *Int. Workshop on Machine Learning for Medical Image Reconstruction*, pp. 112–119 (2018).

Biographies of the authors are not available.