Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Deep learning-based mesoscopic fluorescence molecular tomography: an *in silico* study

Feixiao Long

Deep learning-based mesoscopic fluorescence molecular tomography: an *in silico* study

Feixiao Long*

Beijing QED Technique Co., Ltd., Beijing, China

Abstract. Fluorescence molecular tomography (FMT), as well as mesoscopic FMT (MFMT) is widely employed to investigate molecular level processes *ex vivo* or *in vivo*. However, acquiring depth-localized and less blurry reconstruction still remains challenging, especially when fluorophore (dye) is located within large scattering coefficient media. Herein, a two-stage deep learning-based three-dimensional (3-D) reconstruction algorithm is proposed. The key point for the proposed algorithm is to employ a 3-D convolutional neural network to correctly predict the boundary of reconstructions, leading refined results. Compared with conventional algorithm, *in silico* experiments show that relative volume and absolute centroid error reduce over ~50% whereas intersection over union increases over 15% for most situations. These results preliminarily indicate the promising future of appropriately applying machine learning (deep learning)-based methods in MFMT. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.5.3.036001]

Keywords: mesoscopic fluorescence molecular tomography; deep learning; image reconstruction; *in silico* experiments.

Paper 18069RR received Apr. 4, 2018; accepted for publication Aug. 8, 2018; published online Sep. 4, 2018.

1 Introduction

Fluorescence molecular tomography (FMT) is an imaging modality to retrieve the fluorophore distribution inside the object, which reflects the physiological situations of particular tissues or molecules through measuring the (fluorescence) emission from the surface of media. Among multiple forms of FMT, mesoscopic FMT (MFMT) can be viewed as a special kind of FMT with millimeter-depth differentiation and high resolution.¹ It is reported with proper hardware and algorithm settings; compared with FMT, MFMT can reach 100- to 200- μ m resolution at depths of 3 to 5 mm.² In FMT/MFMT, fluorophore (dye) is mostly used as a probe to monitor molecular or cellular processes.³ For instance, McCann et al.⁴ used FMT combining with magnetic resonance imaging (MRI) to monitor protease activity in mouse tumor. Zacharakis et al.⁵ employed FMT to image a cancer model inside the lungs of nude mice with the help of computed tomography (CT). Muldoon et al.⁶ employed MFMT to analyze the skin lesions. Moreover, FMT/MFMT has been successfully employed in many biomedical applications, such as drug development, cancer detection,⁷ etc., due to its high sensitivity and relatively low cost. However, acquiring accurate depth-localized and less blurry reconstruction still remains challenging, especially with large scattering coefficient turbid media⁸ (due to its nature of leading to more scattered emission light; in fact, any optical imaging modality that utilizes diffusive light will also suffer this issue). One kind of solution to this issue is to employ multimodalities imaging as shown by the most recent applications of FMT/MFMT (also shown above). Multimodalities imaging employs the prior anatomical or structural information of objects obtained from other imaging modalities, such as CT, MRI, or ultrasound imaging, to increase the accuracy of FMT/MFMT reconstructions. Using prior information can indicate the possible locations of fluorophore and thus reduce the measure of

searching space during optimization and increase the robustness of inverse problem (reconstruction) solver.⁹ Yet, this often leads to more complicated device and experimental design. In addition, rigorous coregistration between optical tomography and other imaging modalities, which is generally difficult in most cases, must be performed carefully to get reasonable results. Another conventional technique to improve the performance of depth localization is to try to incorporate the depth information within the inverse problem solver. For example, Tavakoli and Zhu¹⁰ multiplied a diagonal matrix to the sensitivity matrix when performing the reconstruction. The entry of the diagonal matrix is just the maximum singular value of sensitivity matrix at corresponding depth (layer). Shimokawa et al.⁸ employed the hierarchical Bayesian analysis to improve the depth accuracy of reconstruction. Although the above techniques can improve the reconstruction accuracy, in general, it will need extra complex and sometimes time-consuming math operations, such as singular-value decomposition. Herein, it is still incentive to develop a technique to improve the quality of FMT/MFMT reconstruction.

Nowadays, artificial intelligence, especially machine learning (ML)-based methods, have already shown great success in computer vision^{11,12} and natural language processing. Meanwhile, researchers in the medical image community have already begun the exploration for applying ML techniques in medical image segmentation,¹³ histopathological image analysis,¹⁴ and image reconstructions.¹⁵ For example, Sinha et al.¹⁶ employed deep neural network (NN) (U-net) to recover phase objects. Wu et al.¹⁷ employed artificial NN to quickly obtain a fluorescence lifetime image. Yao et al.¹⁵ used a deep NN to directly transfer the detector measurements to fluorescence intensity and lifetime images simultaneously. However, one of the main drawbacks of the above-mentioned research is the lack of 3-D image reconstruction, which limits the possible applications in real preclinical situations. Moreover, the

*Address all correspondence to: Feixiao Long, E-mail: lf832003@gmail.com

biggest problem of applying deep NN as shown above is it obviously covered the clear physical relationship between measurements and reconstructions, thus, turned the reconstruction procedure into a purely data-driven process, which was widely known to be possibly misled by the training dataset,¹⁸ especially when the dataset was flawed or lacked any appropriate physical or mathematical explanations of the trained network.

To address the above-mentioned issues, a two-stage 3-D MFMT reconstruction algorithm is proposed and verified with a series of *in silico* experiments. The method consists of two steps: first, one conventional MFMT reconstruction algorithm (depth-dependent Tikhonov regularization)¹⁹ is employed to acquire preliminary fluorophore distribution; second, a trained 3-D convolutional neural network (CNN) is applied to refine the original fluorescence distribution to get less blurry and depth-localized reconstructions. Here, several points need to be noticed for the proposed two-stage algorithm. First, the application of 3-D CNN is established on the foundation that MFMT reconstruction can be naturally transferred to a 0/1 image segmentation (classification) problem (especially for those applications with constant fluorophore concentration but variable locations, see Sec. 2.2 for more details), the region where ML-based techniques have been successfully and widely employed. Second, although applying ML technique is natural in fluorescence-based optical tomography, the algorithm still preserves the physical essential of MFMT reconstruction through conventional step and only employs CNN in the refine-step to further reduce the original reconstruction error.

The paper is organized as follows. In Sec. 2.1, the optical settings of *in silico* experiments and depth-dependent reconstructions of MFMT will be described. The structure of CNN and the techniques to reduce overfitting will be explained in detail in Sec. 2.2. Meanwhile, the training dataset preparation and the training settings will be given in detail in the same section. Moreover, the quantities to assess the performance of two-stage algorithms are described. In Sec. 3, *in silico* experiments results are shown with explanations. In Secs. 4 and 5, several issues about applying ML-based techniques in FMT/MFMT are discussed; finally, the conclusion is made, respectively.

2 Methods

2.1 Optical Settings and Original Reconstruction

Since the aim of this paper is just to demonstrate the state-of-the-art of proposed algorithms, as well as due to observing the relatively blurry reconstructions in our previous work in the region of MFMT,^{20,21} the phantom with similar size as shown in previous work was used in *in silico* experiments. A cube with dimension $4 \times 4 \times 4 \text{ mm}^3$ was employed as phantom with sources and detectors placed at both the top and bottom surfaces (note the arrangement of sources and detectors is not routine in conventional MFMT; indeed, it imitates some reflective and transmission geometry employed in Refs. 19 and 20). The distance between the adjacent sources (detectors) was set to 1 mm [see Fig. 1(a)] along the x -direction and 0.5 mm along the y -direction, yielding 56 sources and 42 detectors, respectively. The voxel size was set to 0.001 mm^3 , which resulted in 64,000 voxels (40 for each dimension) in the phantom. The absorption coefficient and scattering coefficient were set to 0.3 and 10 mm^{-1} , respectively.¹ The anisotropy and refractive index were set to 0.9 and 1.37, respectively. Note that the set

Fig. 1 (a) Demonstration of sources (red) and detectors (green) locations on the surface of phantom and (b) CNN structures. Note that each convolutional layer also includes batch normalization layer. Fully connected layer contains dropout layer. 0 represents the center voxel is outside the generated objects and 1 means inside.

absorption coefficient is much larger than the one normally employed in FMT.

Since the simulated optical properties and distances between sources and detectors did not satisfy the requirements of applying diffusion approximation to simulate photon propagation in turbid media,¹ Monte Carlo (MC)-based forward–adjoint method was employed to generate sensitivity profiles (matrix). Due to the graphic processing unit (GPU)-based accelerating techniques,²² 10^{11} photons were able to be employed for each source/detector pair in the experiments and the total simulation time was <20 min with the help of four Titan Xp GPUs (Nvidia Corporation, California). Born normalization²³ was employed as shown in the below equation to recover original fluorophore distribution

$$U_B(\mathbf{r}_s, \mathbf{r}_d) = \frac{1}{G^x(\mathbf{r}_s, \mathbf{r}_d)} \int_X \eta(\mathbf{r}) G^x(\mathbf{r}_s, \mathbf{r}) G^m(\mathbf{r}, \mathbf{r}_d) d\mathbf{r}, \quad (1)$$

where $\mathbf{r}_s$ and $\mathbf{r}_d$ represent the spatial coordinate of source and detector, respectively, G^x and G^m are the Green's function of excitation and emission calculated by the MC simulations shown above, respectively, and $\eta(\mathbf{r})$ is the distribution of fluorophore, the quantity to be reconstructed.

The sensitivity profiles were downsampled to half of the original ones for each dimension leading to fewer unknowns (8000 in total) to lower the computational burden of the inverse problem solver. The simulated measurements were acquired through the multiplication of sensitivity matrix and the distribution matrix of fluorophore, which was set to 1 when the matrix representing voxel was located within the region of randomly generated objects in the phantom (see Sec. 2.2.2 for more details) and 0 outside the objects. When performing reconstruction, the following expression is employed:

$$(\mathbf{A}^T \mathbf{A} + \lambda \mathbf{D}) \mathbf{x} = \mathbf{A}^T \mathbf{b}, \quad (2)$$

where $\mathbf{A}$ represents the sensitivity matrix calculated by the forward–adjoint MC method as described above, λ is the Tikhonov

parameter and note that this parameter λ was empirically set to 0.0009 during the experiments, not through the time-consuming method such as L -curve²⁰ in *in silico* experiments, $\mathbf{D}$ is a diagonal matrix whose element is the square root of corresponding entry of diagonal of $\mathbf{A}^T \mathbf{A}$, $\mathbf{x}$ is the fluorophore distribution, and $\mathbf{b}$ denotes the simulated measurement. Readers can refer to Ref. 24 for more details about reconstructions step.

2.2 Refined Reconstruction

2.2.1 Structure of convolutional neural network

After getting the original reconstruction described in Sec. 2.1, a 3-D CNN-based refinement procedure was applied to acquire better depth-localized solution and reduce the blurriness of original reconstructions. It will be natural and reasonable to convert this refined reconstruction problem to a 0 and 1 classification problem (for cases with constant concentration but variable locations), which means using 1 (0) to indicate the location (voxel) where fluorophore exists (does not exist) based on the principles of MFMt.

To achieve this aim, a network with five convolutional layers and two fully connected layers was constructed [see Fig. 1(b)]. In math, all the NNs are actually the compositions of multiple nonlinear real functions. For each convolutional layer, a $3 \times 3 \times 3$ kernel was employed and the stride was set to 1 for three dimensions. In addition, zero padding was filled to each convolutional layer input to make sure the output of the layer was the same as the input. Rectified linear unit¹¹ was employed as the activation function for each neuron except the output layer to mitigate the gradient vanishing or exploding phenomenon during training with the standard back-propagation algorithm. One max pooling layer was inserted between the third and fourth convolutional layer. When employing the above NN to obtain refinement reconstruction, the strategy that decomposed the original reconstruction ($20 \times 20 \times 20$, including whole phantom, Sec. 2.1) into overlapping patches ($7 \times 7 \times 7$ in the experiment, total 2744 patches inside the phantom) was employed [see Fig. 2(a) for demonstration]. In other words, each time one patch was fed into 3-D CNN the network was trained to determine if the center voxel of this patch was inside the boundary of object

(1) or not (0). The above-mentioned procedure was repeated till finishing all the calculations of all patches.

To alleviate the overfitting problem, several techniques were adopted. First, the capacity of the proposed network was limited. The number of neurons for each layer was carefully designed, namely 32 for the first three convolutional layers, 64 for the following two convolutional layers, and 512 for the last two fully connected layers. Second, batch normalization²⁵ technique was employed, which not only accelerates the training process but also plays a role as a regularizer. Third, L_2 regularizer consisting of all weights in the network was added to the loss function. Fourth, dropout layers²⁶ with rate 0.5 were inserted right after every fully connected layer.

Here, one point needs to be noted is that due to the small capacity of proposed network and relative shallow architectures, especially compared with VGG net,²⁷ ResNet,²⁸ or Xception net,²⁹ it is not necessary to employ network structure with skip connection between input and output of a particular convolutional layer as ResNet or densely connected convolutional net,³⁰ which was designed to increase the gradient flow during training originally.

2.2.2 Training dataset and training settings

For simplicity, only balls and ellipsoids that are commonly used as real phantoms¹⁰ were employed in the simulation. The centers of balls and ellipsoids were randomly located inside the center region of the cube ($[0.6, 3.4] \times [0.6, 3.4] \times [0.6, 3.4]$ mm). The radii of balls and lengths of three axes of ellipsoids were uniformly selected from $[0.3, 0.6]$ mm. These settings guarantee that no objects would be beyond the boundaries of the cube. In total, the dataset consisted of 600 random generated balls and ellipsoids, respectively, obeying above criterions. The dataset was then split into two subsets, training and validation dataset with proportion 9:1.

Because the generated objects only occupy small portions of cube, which suggest only small number of training patches will be labeled 1 (positive samples), therefore, a careful selection of negative samples (label 0) must be performed to maintain the balanced proportion of positive and negative samples. Moreover, the chosen negative samples should be representative enough to delineate the situations at the boundary of objects since intuitively the refinement reconstruction is designed to get less blurry and depth-localized reconstruction through correctly predicting the boundary. Thus, in the experiments, most negative samples (90% of total negative samples) were selected from the patches, which at least had some intersections with the object but the patch centers were outside the object boundary.

During training, batch size of 1500 was implemented on single Titan Xp GPU. The number of epochs was set to 600 and the training time for one epoch is <15 s (Ubuntu 16.04, Tensorflow 1.4.0 backend, and Keras 2.1.2). The RMSprop optimizer was employed during training with all the parameters of it default (learning rate 0.001). The L_2 regularization parameter was set to 0.01. After training, the accuracy was over 0.98 for both training and validation datasets.

2.3 Algorithm Verification

2.3.1 Test dataset

One hundred fifty objects (balls and ellipsoids) that obey the same distribution laws shown above were generated separately

Fig. 2 (a) Demonstration of workflow of CNN. The blue rectangle represents the phantom (cube). The green dashed rectangle represents one patch that will be fed into NN. The red ball represents the fluorophore. (b) One slice comparison of ground truth, conventional, and two-stage reconstructions.

and two-stage reconstruction was processed. The accuracy of algorithm was evaluated as follows. First, six statistical quantities, widely employed in classification problem, were adopted, including true positive rate (TPR), true negative rate (TNR), positive predictive value (PPV), false positive rate (FPR), false negative rate (FNR), and accuracy (ACC), to evaluate the error between center voxel predictions and true values for all patches (2744×150). Second, intersection over union (IOU) to access the similarity between the real object and its reconstruction was employed, defined as

$$D = \frac{|R \cap T|}{|R \cup T|},$$

where R and T represent the reconstruction (whatever from conventional or two-stage algorithm) and real object, respectively. When extracting object from conventional algorithm reconstruction, iso-surface with iso-value 0.5 was employed as the boundary. Moreover, the relative volume error and absolute centroid error²⁰ were first calculated for each simulated object. Then, the average values of the above error for all objects were evaluated and shown in Table 2.

2.3.2 Three-dimensional rectangle objects dataset

To test the generality of algorithm preliminarily, 150 3-D rectangles with their centers and edge lengths obeying the same probability laws shown in Sec. 2.2.2 were created in the cube. Again, the quantities discussed in Sec. 2.3.1 were calculated.

2.3.3 Two-objects phantom

It is known that the interparameter cross talk³¹ can affect the quality of reconstructions. Herein, two different objects at different depths were created to test the algorithm performance. In this part, one ball was located at (1.2, 1.4, 1.2) mm with radius 0.3 mm; meanwhile, an ellipsoid whose center was set at (2.0, 2.4, 2.6) mm with three axes length 0.4, 0.3, and 0.5 mm was created, respectively. Similarly, the accuracy of reconstruction was validated using the quantities discussed in Sec. 2.3.1.

Table 1 Statistical quantities to assess the accuracy of algorithm.

	TPR	TNR	PPV	FPR	FNR	ACC
Test dataset	0.960	0.998	0.910	0.002	0.040	0.998
3-D rectangle	0.943	0.998	0.858	0.002	0.057	0.997
Two objects	0.852	0.998	0.902	0.002	0.148	0.995

3 Results

3.1 Test Dataset

The statistical quantities of test dataset are listed in Table 1 first row. From Table 1, the TPR, TNR, and ACC are over 0.96. Meanwhile, the FPR and FNR are <0.05 and PPV is >0.90 . Further, the comparison between the conventional reconstruction and two-stage reconstruction is listed in Table 2 first row. It is clear that the IOU increases over 15% of original one. The relative volume error and absolute centroid error both approximately reduced to their half of conventional reconstruction. To the refinement part, based on current settings, one reconstruction only needs to process $2744 \times 7 \times 7 \times 7$ patches, requiring another 0.22 s in average with single Titan Xp GPU, which is a negligible overhead to total reconstruction time. The fast speed was due to the relative shallow CNN structure as well as the small number of patches (large voxel size and small phantom volume). Herein, it is anticipated with increasing number of patches, longer calculation time will be required.

3.2 Three-Dimensional Rectangle Objects Dataset

It is expected that some statistical quantities are worse compared with the corresponding ones in Sec. 3.1 since 3-D rectangle-related features were not learned by NN during training (Tables 1 and 2 second row). However, the IOU still increases over 15% and relative volume error reduces over 50%, which suggests a two-stage algorithm can still partially correctly predict the distribution of fluorophore. It is reasonable that, although the shapes of 3-D rectangles and balls are quite different, the nonlocalized feature of conventional MFMT reconstruction algorithm⁸ generates blurry and similar reconstructions especially at the surfaces of comparable size of 3-D rectangles and balls. Herein, this phenomenon will still make the network learn some features of reconstructions of 3-D rectangles. Figure 2(b) shows the zx slice ($y = 1.6$ mm) of ground truth, conventional reconstruction, and two-stage reconstruction of one rectangle sample.

Figure 3 shows the merged reconstructions (ParaView, 5.4.1) of another example that there is no difference between the two-stage algorithm and ground truth (IOU = 1). In contrast, the IOU of the conventional reconstruction and ground truth is only 0.77. The white color represents the conventional reconstruction. The green color with different opacity setting represents both the ground truth and two-stage reconstruction due to their complete overlap. In Fig. 3, the light green color shows the overlapping region of conventional reconstruction and two-stage one. Note the white portion that goes beyond the boundary of green rectangle in Fig. 3 [clearly indicated in Figs. 3(a) and 3(b)]. Also in Fig. 3(a), at the bottom of

Table 2 Reconstruction-based parameters to assess the accuracy of algorithm.

	IOU		Volume error %		Centroid error (mm)	
	Conventional	Two stage	Conventional	Two stage	Conventional	Two stage
Test dataset	0.73	0.87	25.0	10.6	0.022	0.011
3-D rectangle	0.69	0.82	50.7	17.9	0.014	0.027
Two objects	0.70	0.78	7.4	5.6	0.084	0.040

Fig. 3 Three projection views of merged reconstruction: (a) yOz , (b) zOx , and (c) xOy .

3-D rectangle (around $z = 1$ mm), it is obvious that the conventional reconstruction was not complete. This phenomenon can also be seen in Fig. 3(b) at the same position. Figure 4 shows the comparison of 3-D view (3-D slicer, 4.6.2) of two-stage reconstruction and conventional reconstruction, respectively. From Fig. 4(a), it can be easily seen less blurry and depth-localized reconstruction of proposed algorithm. Note the irregular parts in Fig. 4(b). Figures 2(b), 3, and 4 visually show that through refinement step depth-localized and less blurry reconstruction was obtained. Only the absolute centroid error of a two-stage algorithm increases compared with the conventional reconstruction.

3.3 Two-Objects Phantom

The statistical quantities are also worse than the results shown in Secs. 3.1 and 3.2 due to the interparameter cross talk. Yet, the IOU, relative volume error, and absolute centroid error all become better compared with conventional reconstruction.

Note that the absolute centroid error was taken from the larger one of two objects centroid error.

4 Discussions

In this paper, the state-of-the-art of two-stage reconstruction was verified through a series of *in silico* experiments. Compared with a conventional reconstruction algorithm, the proposed two-stage reconstruction reduces the relative volume and absolute centroid error dramatically for most situations. Meanwhile, it increases IOU over 15% for all cases. These results preliminarily suggest that through two-stage algorithm, more depth-localized and less blurry reconstructions inside turbid media were obtained. Note that the Tikhonov regularization parameter was not the optimal one in experiments. Herein, it is anticipated with the optimal Tikhonov regularization parameter, lower error and higher IOU of refined reconstructions will be obtained due to optimal conventional reconstructions. Yet, this will also lose the advantage of using two-stage reconstruction since normally finding the optimal Tikhonov regularization parameter is time

Fig. 4 (a) Two-stage algorithm and (b) conventional reconstruction (Video 1, MPEG, 7.0 MB [URL: <https://doi.org/10.1117/1.JMI.5.3.036001.1>]; Video 2, MPEG, 7.0 MB [URL: <https://doi.org/10.1117/1.JMI.5.3.036001.2>]).

consuming. Moreover, if more accurate and reliable conventional MFMT reconstruction algorithm is employed in the first step, it is still expected that overall MFMT performance will be further improved using two-stage algorithm. Due to the nature of diffusive light employed in FMT/MFMT, great effort must be exercised^{8,32} to get less blurry image using conventional methods. In contrast, the proposed one needs less effort to achieve large performances escalation due to the application of ML techniques. However, several points still need to be paid attention to when employing ML technique in FMT/MFMT.

First is that the conventional reconstruction algorithm can be any popular MFMT reconstruction method besides the one employed in this paper (also mentioned above). Three-dimensional CNN is employed in the manner expecting to refine the fluorophore distribution based on the conventional reconstruction not to replace the conventional one. Second, it will be anticipated that the network trained with dataset consisting of more diverse situations, such as measurements with noise, more shapes and sizes of objects, different Tikhonov regularization parameters, or even incorporation with device parameters in real situations etc., would make prediction more practical, although more computational resources are required. Last, I would emphasize that extreme caution must be exercised when using ML-based approaches in FMT/MFMT due to their powerful capacity (multilayer NN can approximate arbitrary function from a mathematical point of view). After all, ML methods are more appropriate for dealing with the problems that do not have definite rules to solve (e.g., to discriminate cats from a picture), not the ones with clear physical explanations as FMT/MFMT.

5 Conclusions

To sum, preliminary *in silico* simulation results show a promising future to employ an ML-based technique to improve the quality of MFMT reconstruction. From the aspect of deep learning technique, one possible future improvement is to explore if a more complex NN structure can further increase the performance of two-stage reconstruction. In addition, the influence of algorithm performance with different size patches can also be investigated in the future. To MFMT, the accuracy and effectiveness of the two-stage algorithm need to be rigorously verified with *ex vivo* or real phantom experiments, which is just the next step plan.

Disclosures

The author declares that there are no conflicts of interest related to this article.

Acknowledgments

The author would like to thank the hardware support of Beijing QED Technique Co., Ltd.

References

1. E. M. C. Hillman et al., "Laminar optical tomography: demonstration of millimeter-scale depth-resolved imaging in turbid media," *Opt. Lett.* **29**(14), 1650–1652 (2004).
2. M. S. Ozturk et al., "Mesoscopic fluorescence molecular tomography of reporter genes in bioprinted thick tissue," *J. Biomed. Opt.* **18**(10), 100501 (2013).
3. F. Stucker, J. Ripoll, and M. Rudin, "Fluorescence molecular tomography: principles and potential for pharmaceutical research," *Pharmaceutics* **3**(2), 229–274 (2011).
4. C. M. McCann et al., "Combined magnetic resonance and fluorescence imaging of the living mouse brain reveals glioma response to chemotherapy," *NeuroImage* **45**(2), 360–369 (2009).
5. G. Zacharakis et al., "Fluorescent protein tomography scanner for small animal imaging," *IEEE Trans. Med. Imaging* **24**(7), 878–885 (2005).
6. T. J. Muldoon et al., "Analysis of skin lesions using laminar optical tomography," *Biomed. Opt. Express* **3**(7), 1701–1712 (2012).
7. K. Wang et al., "Fluorescence molecular tomography in the second near-infrared window," *Opt. Express* **23**(10), 12669–12679 (2015).
8. T. Shimokawa et al., "Hierarchical Bayesian estimation improves depth accuracy and spatial resolution of diffuse optical tomography," *Opt. Express* **20**(18), 20427–20446 (2012).
9. L. Wu et al., "Shape-parameterized diffuse optical tomography holds promise for sensitivity enhancement of fluorescence molecular tomography," *Biomed. Opt. Express* **5**(10), 3640–3659 (2014).
10. B. Tavakoli and Q. Zhu, "Depth-correction algorithm that improves optical quantification of large breast lesions imaged by diffuse optical tomography," *J. Biomed. Opt.* **16**(5), 056002 (2011).
11. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. of the 25th Int. Conf. on Neural Information Processing Systems*, Vol. 1, pp. 1097–1105, Curran Associates Inc. (2012).
12. S. Ren et al., "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1137–1149 (2017).
13. Z. Akkus et al., "Deep learning for brain MRI segmentation: state of the art and future directions," *J. Digit. Imaging* **30**(4), 449–459 (2017).
14. F. A. Spanhol et al., "Breast cancer histopathological image classification using convolutional neural networks," in *Int. Joint Conf. on Neural Networks (IJCNN)*, pp. 2560–2567 (2016).
15. R. Yao et al., "Deep compressive macroscopic fluorescence lifetime imaging," in *IEEE 15th Int. Symp. on Biomedical Imaging (ISBI)* (2018).
16. A. Sinha et al., "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117–1125 (2017).
17. G. Wu et al., "Artificial neural network approaches for fluorescence lifetime imaging techniques," *Opt. Lett.* **41**(11), 2561–2564 (2016).
18. L. Oakden-Rayner, "Exploring the ChestXray14 dataset: problems," <https://lukeoakdenrayner.wordpress.com/2017/12/18/the-chestxray14-dataset-problems> (21 March 2018).
19. R. Endoh, M. Fujii, and K. Nakayama, "Depth-adaptive regularized reconstruction for reflection diffuse optical tomography," *Opt. Rev.* **15**(1), 51–56 (2008).
20. F. Long and X. Intes, "Dental optical tomography with upconversion nanoparticles—a feasibility study," *J. Biomed. Opt.* **22**(6), 066001 (2017).
21. F. Long et al., "Dental imaging using mesoscopic fluorescence molecular tomography: an ex vivo feasibility study," *Photonics* **1**(4), 488–502 (2014).
22. Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Opt. Express* **17**(22), 20178 (2009).
23. V. Ntziachristos and R. Weissleder, "Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized Born approximation," *Opt. Lett.* **26**(12), 893 (2001).
24. F. Long et al., "Dental imaging using laminar optical tomography and micro CT," *Proc. SPIE* **8937**, 89370P (2014).
25. S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proc. of the 32nd Int. Conf. on Machine Learning*, Vol. 37, pp. 448–456, JMLR, Lille, France (2015).
26. N. Srivastava et al., "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
27. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556 (2014).
28. K. He et al., "Deep residual learning for image recognition," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778 (2016).
29. F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1800–1807 (2017).

30. G. Huang et al., "Densely connected convolutional networks," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2261–2269 (2017).
31. S. S. Hou, B. J. Bacskaï, and A. T. N. Kumar, "Optimal estimator for tomographic fluorescence lifetime multiplexing," *Opt. Lett.* **41**(7), 1352–1355 (2016).
32. F. Tian et al., "Algorithmic depth compensation improves quantification and noise suppression in functional diffuse optical tomography," *Biomed. Opt. Express* **1**(2), 441–452 (2010).

Feixiao Long is a medical image processing algorithm engineer at Beijing QED Technique Co., Ltd. He received his PhD degree from Rensselaer Polytechnic Institute in 2016. He received his BE degree in biomedical engineering from Beijing University of Technology in 2006. And he received his ME degree in biomedical engineering from Tsinghua University in 2009.