Multi-attention prior based residual encoder-decoder network for fast and accurate reconstruction in fluorescence molecular tomography

Peng Zhang, Fan Song, Chenbin Ma, Zeyu Liu, Guanglei Zhang*
Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China

ABSTRACT

Fluorescence molecular tomography (FMT) is a powerful modality for resolving the three-dimensional (3D) distribution of fluorescent targets inside biological tissues. However, the inverse problem of the FMT is severely ill-posed due to the strong scattering effects of photons inside biological tissues. Previously, regularization-based methods have been widely used to mitigate the ill-posedness of FMT. Due to the complex iterative computation and time-consuming reconstruction process, the FMT remains an intractable challenge for achieving accurate and fast 3D reconstructions. In this work, we propose a multi-attention prior based residual encoder-decoder network (MAP-REDN) to perform FMT reconstruction. Firstly, the multi-attention mechanism can provide weighted *a priori* information to the fluorescence source, enabling MAP-REDN to effectively mitigate the ill-posedness and enhance the reconstruction accuracy. Secondly, since the direct reconstruction strategy is adopted, the complex iterative computation process in the traditional regularization-based algorithms can be avoided, thus tremendously accelerating the reconstruction process. The experimental results demonstrate the feasibility of the MAP-REDN in achieving accurate and fast FMT reconstruction.

Keywords: Fluorescence tomography, multi-attention, image reconstruction

1. INTRODUCTION

Fluorescence molecular imaging (FMI) can label specific molecules or cells with targeted fluorescent molecular probes to detect physiological and pathological changes *in vivo* at the molecular or cellular levels in advance[1,2]. However, FMI cannot quantitatively detect concentration changes of fluorescent molecular probes. Therefore, fluorescence molecular tomography (FMT) technology has been further presented. Owing to its high sensitivity and specificity, FMT has great application potential in tumor diagnosis, drug development, and treatment evaluation owing to its high sensitivity and specificity[3-5].

However, achieving accurate FMT reconstruction is a challenging task, which comes from the following reasons. Firstly, the simplified photon propagation model is difficult to accurately describe the propagation process of photons in biological tissues, which hinders the improvement of detection accuracy. Secondly, the inverse problem of FMT is severely ill-posed due to the strong scattering of photons in biological tissues. In addition, the inverse problem of FMT is also ill-conditioned because the measured values of the tissue surface are much smaller than the internal nodes, which makes the solution process vulnerable to system noise and model errors, resulting in poor noise immunity and low robustness.

In order to address the severe ill-posedness and ill-conditioned of the inverse problem to a large extent and enhance the reconstruction accuracy, various reconstruction algorithms have been presented in recent decades. On the one hand, feasible region methods[6-8] are proposed, i.e., the interior distribution of the fluorescent source is inferred from its distribution area on the surface of the object, thus reducing the number of unknowns in the equation to be solved. However, due to the subjective nature of the selection of feasible regions, this approach does not fully address the ill-posedness of the inverse problem. On the other hand, more constraints are added to introduce a priori knowledge to the reconstruction process. The constraints are usually implemented by introducing a regularization term into the objective function[9-13]. The FMT reconstruction is converted into an objective function optimization problem with constraints, and the optimal solution of the reconstructed image is obtained through the optimization method. The commonly used penalty term includes the L2 norm[9], L1 norm[10], L0 norm[11], TV[12], group sparsity[13], etc. These methods are usually effective only under certain assumptions; however, these methods typically involve complex iterative computations and time-consuming

* guangleizhang@buaa.edu.cn

Third International Conference on Computer Science and Communication Technology (ICCSCT 2022)
edited by Yingfa Lu, Changbo Cheng, Proc. of SPIE Vol. 12506, 125063F
 · 0277-786X · doi: 10.1117/12.2662505

reconstruction processes. Therefore, it is crucial to explore more efficient methods or solutions to achieve accurate FMT reconstruction.

Recently, deep learning techniques with powerful generalization capabilities have shown great advantages in FMT reconstruction and have obtained more competitive reconstruction performance than traditional algorithms[14-16]. Inspired by these works, we propose a novel multi-attention prior based residual encoder-decoder network for fast and accurate FMT reconstruction. First, the multi-attention mechanism can provide weighted a priori information to the fluorescence source, enabling MAP-REDN to effectively mitigate the ill-posedness to enhance the reconstruction accuracy. Second, since the direct reconstruction strategy is adopted, the complex iterative computation process in the traditional regularization-based algorithms can be avoided more effectively, thus greatly accelerating the reconstruction process. Experiments demonstrate that the proposed MAP-REDN method achieves faster and more accurate reconstructions than the cutting-edge methods.

2. METHOD

2.1 Light propagation model

The light propagation process can be represented by the coupling of diffuse equation (DE) and Robin-type boundary conditions[17,18], which is defined by,

$$\begin{cases} \nabla\cdot[D_x(r)\nabla\Phi_x(r)]-\mu_{ax}(r)\Phi_x(r)=-\delta(r-r_s) & r\in\Omega \\ \nabla\cdot[D_m(r)\nabla\Phi_m(r)]-\mu_{am}(r)\Phi_m(r)=-\Phi_x(r)\eta\mu_{af}(r) & r\in\Omega \\ \Phi_{x,m}(r)+2\rho D_{x,m}(r)\partial\Phi_{x,m}(r)/\partial n=0 & r\in\partial\Omega \end{cases} \tag{1}$$

The above-detailed explanation can be found in References[13,15]. Solving equation (1) with the finite element method (FEM) [19] in the discrete domain yields the following linear relationship,

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

where W is the system matrix, y is the boundary fluorescence measurement, and x denotes the distribution of fluorescent sources to be reconstructed in biological tissues.

Therefore, the FMT inverse problem is to reconstruct the fluorescence distribution x in equation (3). However, the inverse problem is ill-posed owing to the strong scattering effect and insufficient measurement data (i.e., $y<<x$). To alleviate the ill-posedness and obtain good reconstruction results in a robust manner, a regularization term is generally introduced as a constraint term to optimize the inverse problem, which is formulated as equation (3),

$$\arg\min_{x\geq 0}\frac{1}{2}\left\|y-Wx\right\|_2^2+\lambda\left\|x\right\|_\beta \qquad s.t.\ 0\leq\beta\leq 2 \tag{3}$$

where λ denotes the regularization parameter used to balance the regularity term and data term.

2.2 MAP-REDN for FMT

In this section, we develop an deep learning method to establish a mapping relationship between surface photon measurements (y) and internal fluorescence source distribution (x), as shown in equation (4),

$$y=f(x) \tag{4}$$

where f: R → R denotes the deep neural network, y denotes the surface fluorescence measurement signals.

The model structure of the MAP-REDN is shown in Figure 1. The designed multi-attention module is that the feature representations extracted from the network encoder are provided to the decoder module. We define the feature map extracted from the network encoder as $\mathrm{X}_i\in\mathbb{R}^{D\times C_i\times H_i\times W_i}$, where D denotes the depth (i.e., the number of FMT projections), C represents the number of feature maps, and H, W represents the height and width of the FMT projections or feature maps, respectively. Feature compression and fusion are calculated as follows,

$$E_5=W_{c5}^{T}Z_5\oplus Z_4 \tag{5}$$

$$E_4 = (W_{c4}^T (W_4^T E_5)) \oplus Z_3 \tag{6}$$

where $E_5 \in \mathbb{R}^{D\times C_4\times H_4\times W_4}$ represents the fused features of Z_5 and Z_4, $E_4 \in \mathbb{R}^{D\times C_3\times H_3\times W_3}$ represents the fused features of Z_4 and Z_3, $W_{c5} \in \mathbb{R}^{D\times C_5\times C_4}$ and $W_{c4} \in \mathbb{R}^{D\times C_4\times C_3}$ represent the corresponding compressed convolutions, $W_4 \in \mathbb{R}^{D\times C_4\times C_4}$ represents the fused convolution of Z_5 and Z_4, and $\oplus$ represents the feature concatenation. The final output of the multi-branch attention prior module is calculated as follows:

$$Y = W_{out}^T (W_2^T Z_3) \tag{7}$$

where $W_3 \in \mathbb{R}^{D\times C_2\times C_2}$ represents the fused convolution of Z_4 and Z_3, $W_{out} \in \mathbb{R}^{D\times C_3\times 1}$ represents the fused output convolution.

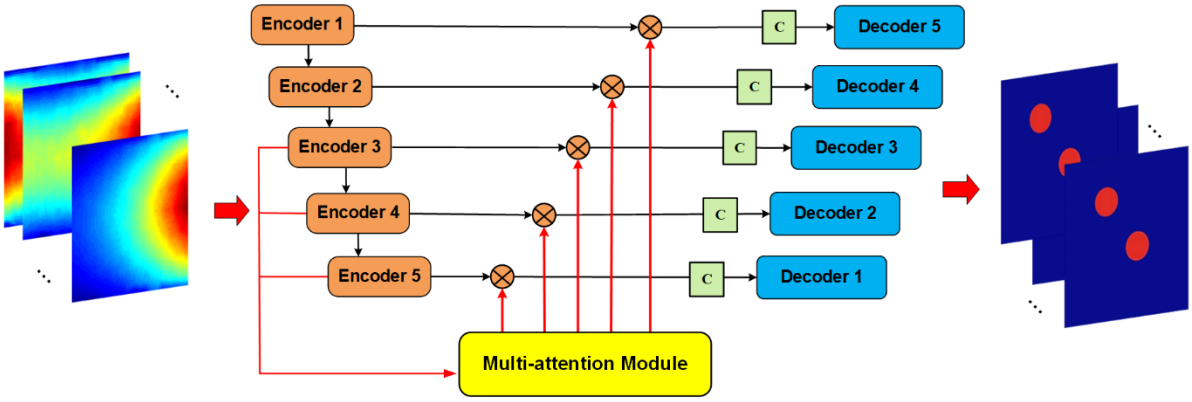

Figure 1. The model structure of the proposed MAP-REDN.

2.3 Image quality evaluation

In this work, we use position error (PE), normalized mean square error (NMSE) and Dice to quantitatively assess the reconstruction performance of the MAP-REDN method,

$$\mathrm{PE} = \left\| P_{rec} - P_{true} \right\|_2 \tag{8}$$

$$\mathrm{NMSE} = \frac{\left\| \hat{\mathbf{x}} - \mathbf{x} \right\|_2^2}{\left\| \mathbf{x} \right\|_2^2} \tag{9}$$

$$\mathrm{Dice} = \frac{2\left| \mathbf{x} - \hat{\mathbf{x}} \right|}{\left| \mathbf{x} \right| + \left| \hat{\mathbf{x}} \right|} \tag{10}$$

here P_{rec} and P_{true} represent the reconstructed tumor centers and ground truth centers, respectively, $\mathbf{x}$ and $\hat{\mathbf{x}}$ represent the actual and reconstructed values, respectively, and $|\mathbf{x}|$ represents the number of voxels in the region $\mathbf{x}$. The larger the Dice, the better the reconstruction performance, with 1 as the upper limit; while the smaller the PE and NMSE, the better the localization performance, with 0 as the lower limit.

3. EXPERIMENTS AND RESULTS

To comprehensively assess the reconstruction capacity of MAP-REND in terms of accuracy, numerical simulation experiments were performed. The classical iteration-based L2-regularized, L1-regularized, and learning-based 3D-En-Decoder methods were utilized for comparison. The reconstruction results of the four different methods are shown in Figure 2.

Figure 2a displays the configurations of the numerical simulation experiments. A cylinder phantom model was put on a rotating stage, the axis of rotation is defined as the Z axis, and the bottom plane is set to Z=0 cm. A excitation light source is placed at a height of Z=0.75 cm to excite the fluorophore to be reconstructed. The fluorescence projection images were

collected at 15° intervals for a full 360° view (i.e., 24 projection data were used for reconstruction). The cylindrical phantom had a height of 1.5 cm and a diameter of 3 cm. Two fluorescent targets (Figure 2a) with a diameter of 3 mm, a height of 5 mm and an edge-to-edge (EED) of 5 mm (Figure 2b) were put inside the cylinder phantom (Figure 2b). Furthermore, a level of 5% Gaussian noise was inserted into the measurement data to better mimic the real experimental situations.

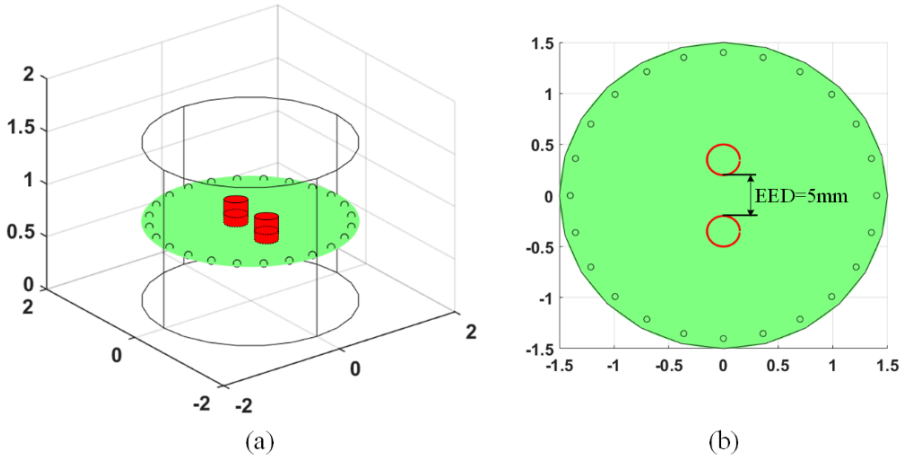

Figure 2. Geometry configuration of the numerical simulations. (a): The 3D display of the cylindrical phantom; (b): The 2D display of the cylindrical fluorfscent targets with the EED of 5mm.

Qualitative analysis is shown in Figures 3a-3d, including a 3D view of the entire imaged phantom and a sliced 2D view of the target localization plane. It can be seen that the L2-regularized method obtains a over-smoothed reconstruction result, which makes the reconstructed fluorescent target much larger than the real target with lower spatial resolution. Second, although the L1-regularized and 3D-En-Decoder methods can approximately reconstruct the tumor region, the two methods face the reconstruction problem of small discontinuous tumors located near the ground truth region. In contrast, the proposed MAP-REDN method achieves high morphological similarity and low position error. Furthermore, a series of evaluation metrics are used to quantitatively evaluate the reconstruction performance. As shown in Figure 3e, compared with the L2-regularized, L1-regularized and 3D-En-Decoder methods, the MAP-REDN has lower PE and NMSE and higher Dice. Therefore, quantitative analysis showed that MAP-REDN performed better in tumor localization and morphological similarity. Therefore, it can be concluded that the multi-attention mechanism is able to provide weighted prior information for the fluorescence sources, alleviating the severe ill-posed nature of the inverse problem and further improving the reconstruction performance.

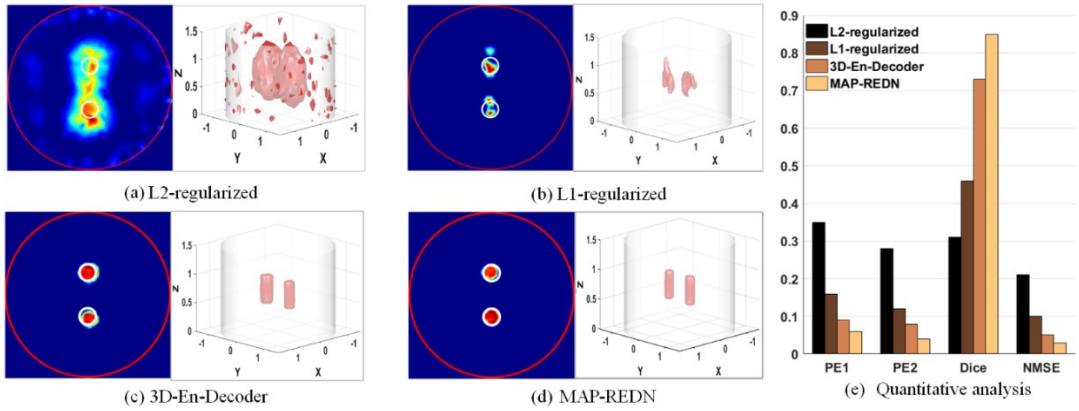

Figure 3. The qualitative and quantitative analysis of the four different reconstruction methods. (a)-(d): The white dotted circles represent the true fluorescent sources. The red polygons clustered around the red spheres in the 3D view are assumed to be reconstructed tumors; (e): The quantitative analysis results of the four different reconstruction methods.

4. CONCLUSION AND DISCUSSION

In this work, we develop a novel MAP-REDN method to achieve fast and accurate FMT reconstruction. The motivation behind this method is that spatial feature mapping and reorganization can be enhanced by introducing a multi-attention mechanism to alleviate the severe ill-conditionedness and enhance the reconstruction accuracy. Experiments show that compared with the traditional L2-regularized, L1-regularized methods and the deep learning-based 3D-En-Decoder method, the MAP-REDN method has improved reconstruction accuracy, morphological similarity and spatial resolution. This method has great potential to provide a potential solution for FMT and alternative imaging models in pursuit of higher reconstruction accuracy.

ACKNOWLEDGMENTS

This work was partially supported by the National Key Research and Development Program of China (No. 2017YFA0700401), the 111 Project (No. B13003), the National Natural Science Foundation of China (No. 61871022), the Beijing Natural Science Foundation (No. 7202102), the Fundamental Research Funds for Central Universities, and the Academic Excellence Foundation of BUAA for PhD Students.

REFERENCES

[1] Ntziachristos, V., "Fluorescence molecular imaging," Annu. Rev. Biomed. Eng. 8, 1-33 (2006).
[2] Graves, E. E., Weissleder, R. and Ntziachristos, V., "Fluorescence molecular imaging of small animal tumor models," Curr. Mol. Med. 4(4), 419-430 (2004).
[3] Rudin, M. and Weissleder, R., "Molecular imaging in drug discovery and development," Nat Rev Drug Discov. 2(2), 123-131 (2003).
[4] Ale, A., Ermolayev, V., Herzog, E., et al., "FMT-XCT: In vivo animal studies with hybrid fluorescence molecular tomography—X-ray computed tomography," Nat. Methods 9(6), 615-620 (2012).
[5] Ntziachristos, V., Tung, C. H., Bremer, C. and Weissleder, R., "Fluorescence molecular tomography resolves protease activity in vivo," Nat. Med. 8(7), 757-761 (2002).
[6] An, Y., Liu, J., Zhang, G., Ye, J., Du, Y., Mao, Y., Chi, C. and Tian, J., "A novel region reconstruction method for fluorescence molecular tomography," IEEE. Trans. Biomed. Eng. 62(7), 1818-1826 (2015).
[7] He, X., Wang, X., Zhang, H., Yi, H. and Hou, Y., "Limited-projection fluorescence molecular tomography based on smoothed l0 norm and feasible region," Chinese Journal of Lasers 45(9), 0907001 (2018).
[8] Cao, X., Zhang, B., Wang, X., Liu, F., Liu, K., Luo, J. and Bai, J., "An adaptive Tikhonov regularization method for fluorescence molecular tomography," Med. Biol. Eng. Comput. 51(8), 849-858 (2013).
[9] Xie, W., Deng, Y., Wang, K., Yang, X. and Luo, Q., "Reweighted L1 regularization for restraining artifacts in FMT reconstruction images with limited measurements," Opt. Lett. 39(14), 4148-4151 (2014).
[10] Guo, H., Yu, J., He, X., Hou, Y., Fang, D. and Zhang, S., "Improved sparse reconstruction for fluorescence molecular tomography with L 1/2 regularization," Biomed. Opt. Express 6(5), 1648-1664 (2015).
[11] Zhao, L., Yang, H., Cong, W., Wang, G. and Intes, X., "L(p) regularization for early gate fluorescence molecular tomography," Opt. Lett. 39(14), 4156-4159 (2014).
[12] Jiang, S., Liu, J., Zhang, G., An, Y., Meng, H. and Tian, J., "Reconstruction of fluorescence molecular tomography via a fused LASSO method based on group sparsity prior," IEEE. Trans. Biomed. Eng. 66(5), 1361-1371 (2018).
[13] Zhang, P., Ma, C., Song, F., Fan, G., Sun, Y., Feng, Y., Ma, X., Liu, F. and Zhang, G., "A review of advances in imaging methodology in fluorescence molecular tomography," Phys. Med. Bio. 67(10), (2022).
[14] Guo, L., Liu, F., Cai, C., Liu, J. and Zhang, G., "3D deep encoder-decoder network for fluorescence molecular tomography," Opt Lett. 44(8), 1892-1895 (2019).
[15] Zhang, P., Fan, G., Xing, T., Song, F. and Zhang, G., "UHR-DeepFMT: Ultra-high spatial resolution reconstruction of fluorescence molecular tomography based on 3D fusion dual-sampling deep neural network," IEEE. Trans. Med. Imaging. 40(11), 3217-3228 (2021).
[16] Huang, C., Meng, H., Gao, Y., Jiang, S., Wang, K. and Tian, J., "Fast and robust reconstruction method for fluorescence molecular tomography based on deep neural network," Proc. SPIE 10881, Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XVII, 108811K (4 March 2019); https://doi.org/10.1117/12.2508468.
[17] Soubret, A., Ripoll, J. and Ntziachristos, V., "Accuracy of fluorescent tomography in the presence of heterogeneities: Study of the normalized Born ratio," IEEE. Trans. Med. Imaging. 24(10), 1377-1386 (2005).
[18] Arridge, S. R., "Optical tomography in medical imaging," Inverse. Probl. 15(2), 41-93 (1999).

[19] Zhang, G., Pu, H., He, W., Liu, F., Luo, J. and Bai, J., "Bayesian framework based direct reconstruction of fluorescence parametric images," IEEE. Trans. Med. Imaging. 34(6), 1378-1391 (2015).