Fast and Robust Reconstruction Method for Fluorescence Molecular Tomography based on Deep Neural Network

Chao Huang^a, Hui Meng^a, Yuan Gao^a, Shixin Jiang^b, Kun Wang^a, Jie Tian^a, *Fellow*, *IEEE* ^a CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences; Beijing Key Laboratory of Molecular Imaging; University of Chinese Academy of Sciences, Beijing,100190, China; ^b School of Computer and Information Technology, Beijing Jiaotong University, Beijing, 100044, China

ABSTRACT

Fluorescence molecular tomography (FMT) is a promising imaging technique in applications of preclinical research. However, the complexity of radiative transfer equation (RTE) and the ill-poseness of the inverse problem limit the effectiveness of FMT reconstruction. In this research, we proposed a novel Deep Convolutional Neural Network (DCNN), Gated Recurrent Unit (GRU) and Multiple Layer Perception (MLP) based method (DGMM) for FMT reconstruction. Instead of establishing the photon transmission models and solving the inverse problem, the proposed method directly fits the nonlinear relationship between fluorescence intensity at the boundary and fluorescent source in biological tissue. For details, DGMM consists of three stages: In the first stage, the measured optical intensity was encoded into a feature vector by transferring the VGG16 model; In the second stage, we fused all encoded feature vectors into one feature vector by using GRU based network; In the last stage, the fused feature vector was used to reconstruct the fluorescent sources by MLP model. To evaluate the performance of our proposed method, a 3D digital mouse was utilized to generate FMT Monte Carlo simulation samples. In quantitative analysis, the results demonstrated that DGMM method has comparable performance with conventional method in tumor position locating. To the best of our knowledge, this is the first study that employed DCNN based methods for FMT reconstruction, which holds a great potential of improving the reconstruction quality of FMT.

Keywords: Fluorescence molecular tomography, Ill-poseness, Deep convolution neural network, Reconstruction.

1. INTRODUCTION

Fluorescence molecular imaging (FMI) can visualize the occurrence and development process of diseases on molecular and cellular levels [1]. Although, FMI is used to obtain the two-dimensional (2D) surface fluorescence, it cannot provide in-depth and quantitative information of inner light sources. This problem severely limits its application for quantitative studies in oncology. Thus, fluorescence molecular tomography (FMT) technology was proposed to acquire the 3D and quantitative information of the *in vivo* fluorescence distribution [2].

The conventional FMT reconstruction depends on the nonlinear radiative transfer equation (RTE) to model the propagation of photons in biological tissues [3]. To reduce the computational complexity of RTE, the approximation model of RTE such as diffusion equation (DE) [2] and the higher-order simplified spherical harmonics approximation (SP_N) [4] were proposed. With the application of Finite Element Method (FEM) or Finite Difference Method (FDM) [5], the approximation model can be transformed to a linear relationship between the photon density on the object surface and the fluorescent source inside the imaging object. Thus, the conventional FMT reconstruction is the inverse problem of this linear relationship. However, the ill-poseness of the inverse problem limits the quality of FMT reconstruction. Thus, many researchers have focused on this problem and proposed different methods. In order to overcome the underdetermined problem of FMT, the measurement with various angles were used widely in FMT reconstruction. Furthermore, anatomical information from X-ray Computer Tomography (CT) or Magnetic Resonance Imaging (MRI) was utilized to build the approximation model of RTE [6]. Besides, the L_2 -norm regularization (also known as Tikhonov regularization) method was used widely for its convenient implementation [2, 7]. However, the result acquired from L_2 -norm regularization turned out to be over smoothed. In order to tackle the over-smooth problem, sparse prior has been adopted [8, 9]. Based on the sparse prior, sparse regularization (L_0 , L_1 , L_p etc.) [9-12] or greedy strategies have been used to ensure sparsity of solutions [13-15]. Although these methods have gradually improved the reconstruction quality of FMT, they still have not achieved satisfactory results for FMT reconstruction. In the conventional FMT reconstruction, the nonlinear RTE is extensively approximated by DE or SP_N , which inevitably causes errors in FMT reconstruction [12, 16, 17]. This fundamental problem has not been solved well.

In this paper, we propose a novel DCNN [18], GRU [19] and MLP [20] based method (DGMM) for fast and accurate FMT reconstruction. Instead of establishing the simplified photon transmission models and solving the ill-pose inverse problem, DGMM directly fits the nonlinear relationship between the measured fluorescence intensity at the boundary and the fluorescent source in biological tissue. Therefore, DGMM can essentially remove the deviation originated from the simplification of RTE and the ill-poseness of the inverse problem. These properties make it more suitable for fast and accurate FMT reconstruction. Simulation experiments show that our proposed method achieved accurate FMT reconstruction.

This paper is organized as follows: Section II describes the reconstruction algorithm DGMM. Section III presents the results of the simulation experiments. At last, Section IV gives the conclusion and discussion.

2. METHOLOGY

In details, our proposed model DGMM consists of three stages: information encode stage, encoded information fuses stage and reconstruction stage. In the first stage, the measured optical intensities acquired from eight different excitation location were encoded into feature vectors by transferring the VGG16 model [21]:

$$Feature_j = Model_{first-stage}(\Phi_j), j = 1, 2, ..., 8$$
⁽¹⁾

Where Φ_j denotes the density of surface photon and *j* represents the excitation location. *Feature_j* is the inferenced feature vector obtained from *Model_{first-stage}*. The architecture of *Model_{first-stage}*, which contains 13 convolutional layers and 5 max pooling layers is designed using the same principles inspired by VGG16 model. In order to overcome the overfitting problem, the parameters of the former 10 convolutional layers are transferred from the trained VGG16 model and freezed. The filters which have a 3 × 3 receptive field is used; the stride of convolutional layer is fixed to 1 pixel; the spatial padding of convolution layer input is well designed in order to preserve the spatial resolution after convolution. Maxpooling is performed over a 2 × 2 receptive window, with stride 2. The non-linear activation function following the convolution is called rectified linear unit (ReLU). In the second stage, we fused all encoded feature vectors that obtained from the same sample into one feature vector by using GRU recurrent neural network:

$$z_{i} = \sigma (U_{z} Feature_{i} + W_{z} h_{i-1})$$
⁽²⁾

$$r_{j} = \sigma \left(U_{r} Feature_{j} + W_{r} h_{j-1} \right)$$
(3)

$$\widetilde{h}_{j} = tanh\left(UFeature_{j} + W(r_{j} \odot h_{j-1})\right)$$
(4)

$$h_j = o_j = (1 - z_j) \times h_{j-1} + z_j \times \tilde{h_j}$$
⁽⁵⁾

Where z_j denotes the update gate, which decides how much the recurrent unit updates its activation or content; r_j is a set of reset gates, which is used to control the extent to which to forget the previous state; $\tilde{h_j}$ represents a candidate activation calculated similarly to that of the traditional recurrent unit and \odot is an element-wise multiplication operator. h_j , h_{j-1} is unit state and previous unit state respectively; o_j is the output of the recurrent unit. For the sequence *Feature_j*, j = 1, 2, ..., 8the gated recurrent unit perform the same task for every element, with the output depended on the previous computations. And the output o_8 is used in the third stage to reconstruct the fluorescence target. In the last stage, the fused feature vector was used to reconstruct the position and shape of fluorescent sources by multiple layer perception model (MLP). Given that Multi-task Learning can exploit commonalities and differences across tasks and improve the learning efficiency and prediction accuracy, the output of MLP is designed not only to accurately locate the tumor barycenter \hat{P} but also to obtain the shape of fluorescence target \hat{S} . The tumor barycenter \hat{P} is acquired from MLP, with two hidden layers that have 1024 and 64 channels respectively; the reconstructed fluorescence source \hat{S} is calculated directly from the output of gated recurrent unit (o_8). Rectified linear unit (ReLU) is used as the non-linear activation function. In the hidden layers, dropout (probability = 10%) is utilized to reduce the overfitting problem. The details of our proposed model is given in Fig. 1. In order to train our proposed method, the optimization function is given by:

$$Loss = \|P - \hat{P}\|_{1} + \lambda CrossEntropyLoss(S, \hat{S})$$
(6)

$$CrossEntropyLoss(S, \hat{S}) = -\left(Slog(\hat{S}) + (1 - S)log(1 - \hat{S})\right)$$
(7)

where *P* represents the true position; *S* denotes the true fluorescence target; λ is the regularization parameter used to balance the influence between the position error and the shape similarity. $\lambda = 10$ has been used in this study.



Figure 1. The architecture and process scheme of our proposed method (DGMM). (a) The information encode stage, the input sequence is consisted by eight RGB images obtained from different excitation location. In the neural network, the grey block denotes traditional convolutional layers and the red block represents the max-pooling layers. Then the encoded information *Feature_j* is inferenced by the neural network corresponding to the input Φ_j . (b) The encoded information fuse stage. (c) The reconstruction stage.

3. EXPERIMENT AND RESULT

In order to train the connection parameters and validate our proposed method DGMM, we first created the sample set. The sample set contains both surface photon density distribution Φ_j and true fluorescence target *P* and *S*. Obtaining a large sample set by gathering a huge number of real mouse models is extremely unpractical, which can be the major challenge of applying DGMM. Since Monte Carlo (MC) method can simulate the photon transport in biological tissues and obtain accurate and reliable results [22, 23], we adopted MC method to create the sample set for DGMM. Then, 3000 cases of FMT simulations with 375 types of single fluorescence source were collected to build the single source sample set by using MC method and MOSE 2.3 [23].

The simulation mouse model (46857 nodes and 258557 elements) used in MC method was acquired from a segmented Xray CT mouse imaging, which contained five different organs: liver, heart, lung, kidney and muscle. The corresponding optical absorption and scattering coefficients of different organs are listed in Table 1 [15]. The simulation mouse model is shown in Fig. 2(a). A spherical fluorescence sources S with 1.0mm diameter was placed in the liver. The white dots in Fig. 2(b) represent the location of the different excitation light sources, which were modeled as isotropic point sources located one mean free path beneath the surface in the z = 12.5mm plane. For each excitation location, the emitted surface fluorescence measurement with a 160 degree field of view was obtained from the opposite side of the excitation light source (shown in Fig. 2(b)). Parameters in DGMM were learned by minimizing the loss function mentioned above. The optimizer utilized in the network was stochastic gradient descent (SGD) with learning rate: 0.1, learning rate decay: 0.9 and regularization rate: 0.0001. The DGMM network was implemented by Tensorflow 1.4.0 backend in python 2.7. It was trained with epochs = 50, batch size = 16. All computer process were accomplished by a desktop computer with a 3.40GHz Intel Core i7 CPU, 12GB RAM and GTX1080 Ti GPU.

Table 1. Optical parameters of the heterogeneous model.

Material	$\mu_{ax}(mm^{-1})$	$\mu_{sx}(mm^{-1})$	$\mu_{am}(mm^{-1})$	$\mu_{sm}(mm^{-1})$
Muscle	0.0052	1.08	0.0068	1.03
Lung	0.0133	1.97	0.0203	1.95
Heart	0.0083	1.01	0.0104	0.99
Liver	0.0329	0.70	0.0176	0.65
Kidney	0.0660	2.25	0.0380	2.02



Figure 2. The views of the heterogeneous mouse for numerical simulation. (a), (b) is the 3-D and axial sectional views of the model respectively. One target source S is implanted inside the liver and the center of the source is in the z = 12.5mm plane.

In order to evaluate the performance of our proposed method DGMM, two simulation experiments were designed in this section. To evaluate the reconstruction results quantitatively, the barycenter error (BCE) between the reconstructed source and the true source was computed.

Experiment 1: The positioning accuracy of DGMM in the single-source FMT simulation samples were evaluated by using the five-fold cross validation. The maximum BCE, mean BCE, and the number of test samples whose BCE is greater than 0.3 mm was calculated to evaluate the DGMM performance.

Experiment 2: To evaluate the robustness of network with different tumor depth and organ distribution, a mesh with a larger liver was generated and four single source samples with different source depths were simulated in this mesh. In this experiment, the simulated samples were only used as test samples and the single source sample set was used as the training set. Furthermore, the iterated shrinkage based method with the L1-norm (IS_L1) [10] was used for comparison and the reconstruction results are presented in Fig. 3.

The quantitative analysis of Experiment 1 (Table 2) proved that DGMM achieved accurate positioning of the fluorescence source. The mean and maximum BCE of the single-source reconstruction in five-fold cross validation were only 0.263 mm and 0.332 mm, respectively. Furthermore, there were only 6 (8% in 75 test samples) DGMM reconstructions per 75 tests with BCE > 0.3 mm. These results revealed the feasibility and outstanding accuracy of DGMM in single-source reconstructions. Results of Experiment 2 demonstrated that DGMM still achieved accurate source reconstruction, when the depth of fluorescent target were varied from 1.8mm to 3.6mm (Fig. 3(b)-(e)). The reconstruction results in Fig.3 revealed the good robustness of DGMM for the varied tumor depth and organ distribution.

Table 2. Quantitative evaluation of DGMM for single source reconstruction.

Quantitative Index	Barycenter Error		
Mean	0.263mm		
Maximum	0.332mm		
Quantitative Index	Average Sample Number Per 75 tests		
BCE > 0.3mm	6 (8% in 75 tests)		



Figure 3. Single source FMT reconstruction results in varied tumor depth and organ distribution. (a) The illustration of different organ distribution between the standard and organ varied meshes; (b) - (e) axial slices showing the IS_L1, DGMM reconstructed results and the true sources (black circles). The depths increase from 1.8 mm to 3.6 mm.

4. DISCUSSION AND CONCLUSION

In conclusion, we proposed a novel Deep Convolutional Neural Network (DCNN), Gated Recurrent Unit (GRU) and Multiple Layer Perception (MLP) based method (DGMM) to achieve accurate and fast FMT reconstruction. Instead of establishing the simplified photon transmission models and solving the ill-pose inverse problem, DGMM directly fits the nonlinear relationship between the measured fluorescence intensity at the boundary and the fluorescent source in biological tissue, which makes it different from all the other existing FMT reconstruction strategies. The parameters in DGMM were trained from thousands of simulated single source samples, and the designed simulated single-source experiments revealed that DGMM had comparable reconstruction results with the conventional IS_L1 method in tumor positioning accuracy. However, there are still obvious disadvantages of the proposed DGMM, such as different sample sets need to be established corresponding to different tumor types; it has cannot reconstruct the 3D tumor morphology besides the location; it yet not be utilized in two or more sources FMT reconstruction. Our future work will focus on solving these challenges and we believe that this novel strategy is of great potential to open a new gate for improving *in vivo* optical diffusion tomography.

REFERENCES

- Schoder, H., and Ong, S. C., "Fundamentals of molecular imaging: Rationale and applications with relevance for radiation oncology," Seminars in Nuclear Medicine, 38(2), 119-128 (2008).
- [2] Bangerth, W., and Joshi, A., "Adaptive finite element methods for the solution of inverse problems in optical tomography," Inverse Problems, 24(3), (2008).
- [3] Ntziachristos, V., Ripoll, J., Wang, L. H. V. et al., "Looking and listening to light: the evolution of whole-body photonic imaging," Nature Biotechnology, 23(3), 313-320 (2005).
- [4] Arridge, S. R., "Optical tomography in medical imaging," Inverse Problems, 15(2), R41-R93 (1999).
- [5] An, Y., Liu, J., Zhang, G. L. et al., "A Novel Region Reconstruction Method for Fluorescence Molecular Tomography," Ieee Transactions on Biomedical Engineering, 62(7), 1818-1826 (2015).
- [6] Wang, D. F., Liu, X., Chen, Y. P. et al., "A Novel Finite-Element-Based Algorithm for Fluorescence Molecular Tomography of Heterogeneous Media," Ieee Transactions on Information Technology in Biomedicine, 13(5), 766-773 (2009).
- [7] Wang, D. F., Song, X. L., and Bai, J., "Adaptive-mesh-based algorithm for fluorescence molecular tomography using an analytical solution," Optics Express, 15(15), 9722-9730 (2007).
- [8] Baritaux, J. C., Guerquin-Kern, M., and Unser, M., "Integrated Modeling and Reconstruction with Sparsity Constraints for Fdot," 2009 Ieee International Symposium on Biomedical Imaging: From Nano to Macro, Vols 1 and 2, 173-176 (2009).
- [9] Mohajerani, P., Eftekhar, A. A., Huang, J. D. et al., "Optimal sparse solution for fluorescent diffuse optical tomography: theory and phantom experimental results," Applied Optics, 46(10), 1679-1685 (2007).
- [10] Han, D., Tian, J., Zhu, S. P. et al., "A fast reconstruction algorithm for fluorescence molecular tomography with sparsity regularization," Optics Express, 18(8), 8630-8646 (2010).
- [11] Cao, N., Nehorai, A., and Jacob, M., "Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm," Optics Express, 15(21), 13695-13708 (2007).
- [12] Han, D., Tian, J., Liu, K. et al., "Sparsity-Promoting Tomographic Fluorescence Imaging With Simplified Spherical Harmonics Approximation," Ieee Transactions on Biomedical Engineering, 57(10), 2564-2567 (2010).
- [13] Han, D., Yang, X., Liu, K. et al., "Efficient reconstruction method for L1 regularization in fluorescence molecular tomography," Applied Optics, 49(36), 6930-6937 (2010).
- [14] Ye, J. Z., Chi, C. W., Xue, Z. W. et al., "Fast and robust reconstruction for fluorescence molecular tomography via a sparsity adaptive subspace pursuit method," Biomedical Optics Express, 5(2), 387-406 (2014).
- [15] Ye, J. Z., Du, Y., An, Y. et al., "Reconstruction of fluorescence molecular tomography via a nonmonotone spectral projected gradient pursuit method," Journal of Biomedical Optics, 19(12), (2014).
- [16] Chen, X. L., Zhang, Q. T., Yang, D. F. et al., "Hybrid radiosity-SP3 equation based bioluminescence tomography reconstruction for turbid medium with low- and non-scattering regions," Journal of Applied Physics, 115(2), (2014).
- [17] Yang, D. F., Chen, X. L., Cao, X. et al., "Performance investigation of SP3 and diffusion approximation for threedimensional whole-body optical imaging of small animals," Medical & Biological Engineering & Computing, 53(9), 805-814 (2015).
- [18] Alippi, C., Disabato, S., and Roveri, M., "Moving Convolutional Neural Networks to Embedded Systems: the AlexNet and VGG-16 case," 2018 17th Acm/Ieee International Conference on Information Processing in Sensor Networks (Ipsn), 212-223 (2018).
- [19] Davis, A. S., and Arel, I., "Faster Gated Recurrent Units via Conditional Computation," 2016 15th Ieee International Conference on Machine Learning and Applications (Icmla 2016), 920-924 (2016).
- [20] Kollarits, J., "Concerning the multiple layering and combination of visual perception and its analogy with art therapy, directed dreaming and hallucinations.," Journal Fur Psychologie Und Neurologie, 22, 171-175 (1918).
- [21] Simonyan, K., and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," International Conference on Learning Representations, (2014).
- [22] An, Y., Liu, J., Zhang, G. L. et al., "Compactly Supported Radial Basis Function-Based Meshless Method for Photon Propagation Model of Fluorescence Molecular Tomography," Ieee Transactions on Medical Imaging, 36(2), 366-373 (2017).
- [23] Ren, S. H., Chen, X. L., Wang, H. L. et al., "Molecular Optical Simulation Environment (MOSE): A Platform for the Simulation of Light Propagation in Turbid Media," Plos One, 8(4), (2013).