2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 979-8-3503-2447-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/EMBC40787.2023.10340506

FISTA-NET: Deep Algorithm Unrolling for Cerenkov luminescence tomography

Xin Cao[1], Mengfei Du[1], Yi Chen[1], Gege Zhang[1], Jun Zhang[1], Weitong Li[1], Kang Li[1,*] and Fengjun Zhao[1,*]

***Abstract*— Cerenkov luminescence tomography (CLT) is a highly sensitive and promising imaging technique that can be used to reconstruct the three-dimensional distribution of radioactive probes in living animals. However, the accuracy of CLT reconstruction is limited by the simplified radiative transfer equation and ill-conditioned inverse problem. To address this issue, we propose a model-based deep learning network that combines the neural network with a model-based approach to enhance the performance of CLT reconstruction. The Fast Iterative Shrinkage Thresholding Algorithm (FISTA), a traditional model-based approach, is expanded into a deep network (known as FISTA-NET). Each layer in the network represents an iteration of the algorithm steps, and connecting these layers can form a deep neural network. In addition, different from the traditional FISTA, the key parameters in FISTA, such as gradient step size and threshold value, can be learned through training data without manual production. To evaluate the performance of FISTA-NET, numerical simulation experiments were conducted, which demonstrate its excellent positioning and shape recovery abilities.**

***Clinical Relevance*—This indicates that FISTA-NET strategy can significantly improve the quality of CLT reconstruction, which is further beneficial to the assessment of disease activity and treatment effect based on CLT.**

I. Introduction

When charged ions travel faster than the speed of light through a medium, they emit visible light, which we call Cerenkov radiation [1]. Cerenkov luminescence imaging (CLI) is a highly sensitive imaging technique, which is based on Cerenkov radiation and combines the advantages of optical imaging and radionuclide imaging. In recent years, CLI has been widely used in clinical research, such as human thyroid imaging, Cerenkov luminescence endoscope biomedical imaging, and so on. However, CLI is a planar imaging method and cannot provide the three-dimensional distribution of radioactive probes. Therefore, Cerenkov luminescence tomography (CLT) has been proposed as a 3D optical imaging technology based on CLI [2].

CLT can recover the 3D distribution of radioactive probes in biological tissues from surface optical signals. However, obtaining satisfactory CLT reconstruction results remains challenging due to the ill-posed nature of the inverse problem, which will lead to the high complexity of reconstruction and affect the performance of CLT reconstruction. In order to solve the ill-posed problem, researchers conducted an in-depth study. Many methods have been applied to CLT reconstruction, including model-based methods and deep learning methods.

In the model-based method, the diffusion equation and simplified spherical harmonic equation are used to simplify the radiation transfer equation (RTE) in order to reduce the computational complexity. Many researchers also try to add some prior information to improve the reconstruction performance, including animal structure prior[2, 3], Cerenkov spectrum prior[4, 5], and Cerenkov source distribution prior. Among them, the sparsity of the source is often used as the Cerenkov source distribution prior, which is usually expressed as regularization methods, such as L_2 -norm regularization, L_1 -norm regularization and L_P -norm regularization. Although the above methods improve the reconstruction performance of CLT, model-based methods usually solve inverse problems through iterative optimization, which takes a long time.

In recent years, neural network methods have been used for CLT reconstruction. Different from traditional methods, neural network directly learn the relationship between surface photon intensity and radiation source distribution to carry out CLT reconstruction. This approach can fundamentally avoid approximation errors through end-to-end learning. In order to improve the accuracy and speed of single-source reconstruction, a network based on gated recursive elements is proposed. Meanwhile, Zhang et al. applied multi-layer fully connected neural network to CLT reconstruction, which improved the precision of CLT reconstruction[6]. Cao et al. proposed a stackable denoising autoencoder (SADE) framework for CLT reconstruction[7], which obtained accurate reconstruction results by continuously shrinking the permission region. Meng et al. proposed a local connection network based on K-Nearest Neighbor (KNN) for FMT reconstruction [8], which improved the accuracy and stability of FMT reconstruction. Zhang et al. applied the attentional local connection network to CLT reconstruction, which further improved the accuracy of reconstruction[9]. In addition, Li et al. used 3D graph convolutional neural networks for CLT reconstruction, which takes three-dimensional information into account. Although the

This work was supported in part by the National Key Research and Development Program of China (2019YFC1521102, 2019YFC1521103); Key Research and Development Program of Shaanxi Province (2019GY215, 2021ZDLSF06-04); National Natural Science Foundation of China (61701403,61806164); China Postdoctoral Science Foundation (2018M643719); Graduate Innovation Program of Northwest University (CX2023185).

1 is with the School of Information Sciences and Technology, Northwest University, Xi'an, 710127, China

* is corresponding author (likang@nwu.edu.cn, fjzhao@nwu.edu.cn).

application of deep learning in CLT reconstruction has made some progress, the accuracy of reconstruction is still affected by the excessively sparse data and the differences between data sets, and the neural network method lacks the interpretability compared with the model-based method.

In this study, we combined the traditional model-based approach with the neural network approach and designed an effective deep structure based on the FISTA, which can improve the accuracy of CLT reconstruction and overcome the shortcomings of poor interpretability of the neural network. The method, known as FISTA-NET, works by unrolling FISTA into a deep network consisting of five layers, each corresponding to an iteration of the FISTA. We take the surface optical signal data as input, flow the data over the network, and finally output the reconstructed radiation source distribution. All parameters of the algorithm (gradient step, threshold, etc.) are obtained by deep network learning. Numerical simulation experiment was conducted to evaluate the performance of FISTA-NET, and the FISTA was used as the control group. The experimental results show that our method performs well in the ability of source positioning and shape recovery.

II. METHOD

A. Photon propagation model

In the model-based reconstruction method, the RTE is used to describe the process of light transport in the organism. Because RTE contains multiple variables, it is almost impossible to solve analytically. To address this issue, the simplified diffusion approximation (DA) equation is proposed to describe the optical transport process, which has been widely used in CLT, as well as other optical molecular tomography. Based on finite element analysis, the linear relationship between surface photon intensity and Cerenkov source distribution inside the object can be established as follows：

$$Ax = b \tag{1}$$

where A represents the system matrix, x is the distribution of Cerenkov sources in the organism, and b is the intensity of photons on the organism surface.

B. FISTA

Iterative shrink-thresholding Algorithm (ISTA) is considered to be a time-consuming method. FISTA improves ISTA, which can effectively alleviate ill-posed problems in the process of CLT reconstruction. The objective function of CLT reconstruction can be expressed as:

$$minE(X) = \frac{1}{2}\|Ax - b\|_2^2 + \lambda\|x\|_1 \tag{2}$$

The FISTA solves Eq. (2) by iterating the following steps:

$$\begin{cases} x^{(k)} = \mathcal{T}_\alpha\left(y^{(k)}\right) - \mu A^T(Ay^k - b) \\ t^{(k+1)} = \dfrac{1+\sqrt{1+4(t^{(k)})^2}}{2} \\ y^{(k+1)} = x^k + \left(\dfrac{t^k - 1}{t^{(k+1)}}\right)(x^k - x^{k-1}) \end{cases} \tag{3}$$

The main improvement made by the FISTA over the ISTA was that the iterative contraction operator $\mathcal{T}_\alpha$ was no longer applied directly to $x^{(k)}$ but to $y^{(k)}$, using a carefully designed linear combination of the previous two estimates of $x^{(k)}$, $x^{(k-1)}$. In addition, FISTA only one cleverly selected additional estimation.

C. CLT Reconstruction Based on FISTA-Net

FISTA-NET consists of three modules, including gradient descent module $r^{(k)}$, proximal mapping module $x^{(k)}$, and momentum module $y^{(k)}$. Fig 1. shows the architecture of FISTA-NET.

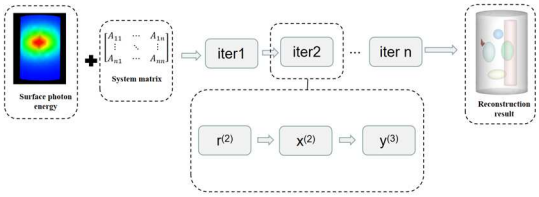

Fig 1. Structure diagram of FISTA-NET.

Gradient descent module $r^{(k)}$. This module updates the Cerenkov source distribution by an equation based gradient descent operation. Given $y^{(k)}$ as the output of the previous layer. Obviously, the objective is to obtain a more precise estimate that minimizes $\|Ax - b\|_2^2$.

$$\begin{cases} r^{(k)} = y^{(k)} - \mu^k A^T\left(Ay^{(k)} - b\right) \\ x^{(k)} = \arg\min_x \dfrac{1}{2}\left\|x - r^{(k)}\right\|_2^2 + \lambda\|x\|_1 \\ y^{(k+1)} = x^{(k)} + \rho^k(x^{(k)} - x^{(k-1)}) \end{cases} \tag{4}$$

Proximal mapping module $x^{(k)}$. The goal of this module is to calculate $x^{(k)}$ from Eq. (4), which has an input of $r^{(k)}$. The solution of $x^{(k)}$can be obtained from Eq. (4) as follows:

$$x^{(k)} = soft\left(r^{(k)}, \theta\right) \tag{5}$$

θ as the shrinkage threshold, is a learnable parameter in this module.

momentum module $y^{(k)}$. This module increases the rate of convergence by introducing the momentum term, as it did in FISTA, which was shown to converge to $O(1/k^2)$ in function values compared to ISTA's slower $O(1/k)$ rate[10]. Its faster convergence rate is due to the clever selection of update weights for the first two results without the need for additional gradient evaluation. In FISTA-NET, we continue this advantage by replacing the constant update weight $t^{(k)}$ with the learnable parameter $\rho^{(k)}$, learned autonomously from the training data set.

D. Model-based parameter constraints

Although parameters such as $\mu^{(k)}$, $\theta^{(k)}$, $\rho^{(k)}$ are dynamically learnable in FISTA-NET, additional constraints are introduced to ensure that they converge correctly. Based on the observation that non-goose-step sizes and thresholds can be generated during iteration. Therefore, a good rule of thumb is to make $\mu^{(k)}$, $\theta^{(k)}$, $\rho^{(k)}$ positive. In addition, the gradient step size $\mu^{(k)}$ should decay smoothly with iteration. The threshold value $\theta^{(k)}$ should also decrease iteratively, and

the two-step update weight $\rho^{(k)}$ should correspond to the two-step monotone increase.

$$\begin{cases} \mu^{(k)} = sp(w_1 k + c_1), \ w_1 < 0 \\ \theta^{(k)} = sp(w_2 k + c_2), \ w_2 < 0 \\ \rho^{(k)} = \dfrac{sp(w_3 k + c_1) - sp(w_3 + c_3)}{sp(w_3 k + c_3)}, \ w_3 > 0 \end{cases} \tag{6}$$

where the softplus function $sp(x) = \ln(1 + \exp(x))$; $\rho^{(k)} \in (0,1)$ is consistent with FISTA; One benefit of the softplus function is its simple derivative function. The parameters $\{w_1, w_2, w_3, c_1, c_2, c_3\}$ are initialized according to experience[11].

III. Experiments Setting

In this part, to evaluate the performance of FISTA-NET, we conducted several sets of numerical simulation experiments, including single-source experiment, dual-source experiment and controlled experiment, in which FISTA was adopted as the baseline.

A. Data collection

For the numerical simulation experiments, we utilized a cylindrical model with a height of 30mm and a radius of 10mm to simulate the biological body. The cylinder model mainly includes five organs: heart, bone, liver, lung and muscle, as shown in Fig 2. (a). The cylinder model is discretized into a unified tetrahedral grid containing 4626 nodes and 25,480 tetrahedral elements, as shown in Fig 2. (b). The optical parameters of each organ are shown in Table I, in which the photon wavelength is set at 650mm.

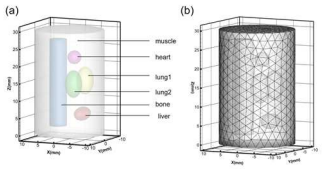

Fig 2. (a) is the three-dimensional view of the cylindrical imitation model. (b) is the standard mesh.

TABLE I. Optical Properties

Tissue	$\mu_a(mm^{-1})$	$\mu_s(mm^{-1})$	g
Muscle	0.016	0.510	0.90
Heart	0.011	1.053	0.86
Bone	0.021	2.864	0.90
Liver	0.065	0.723	0.90
Lung	0.036	2.246	0.90

B. Quantitative evaluation

In order to conduct quantitative evaluation of experimental results, we adopted LE and DICE coefficient as evaluation indexes. LE is the error between the actual position of the radiation source and the reconstructed position, which is calculated as follows:

$$LE = \|L_R - L_T\|_2 \tag{7}$$

where L_R is the central coordinate of the reconstructed radiation source, and L_T is the central coordinate of the real radiation source. $DICE$ coefficient is used to evaluate the accuracy of Cerenkov source shape recovery in CLT reconstruction, and its definition is as follows:

$$DICE = \frac{2|X_r \cap Y_a|}{|X_r| + |Y_a|} \tag{8}$$

where X_r represents the reconstruction area and Y_a represents the actual light source area. The DICE coefficient ranges from 0 to 1, with the value closer to 1 indicating better shape recovery.

TABLE II. Quantitative Results of Single-Source

source	Coordinate of Cerenkov source(mm)	LE	DICE
Model 1	(-6 6 17)	0.405	0.790
Model 2	(-6 -5 18)	0.642	0.674

IV. Results

A. Single-source reconstruction

Single-source numerical simulation experiments were carried out, in which two samples of Model 1 and Model 2 were used to measure the single-source reconstruction results of different methods, whose coordinates were set as (-6, 6, 17) mm and (-6, -5, 18) mm respectively. The experimental results are shown in Fig 3. Secondly, LE and DICE coefficients were calculated, and the quantitative analysis results were shown in Table II. The experimental results show that FISTA-NET has better positioning accuracy and shape recovery ability in single-source environment.

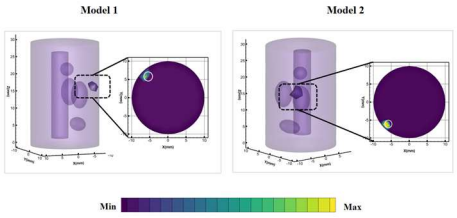

Fig 3. is the 3D view and 2D cross section of the reconstruction results of Model 1 and Model 2 using FISTA-NET method. Axial view is selected as a 2D cross section, and the white circle represents the position of the true Cerenkov source.

B. Dual-source reconstruction

Dual-source numerical simulation experiments were carried out, in which two samples of Model 3 and Model 4 were used to measure the dual-source reconstruction performance of FISTA-NET respectively. The coordinate of Model 3 was set as (-6, -5, 18) mm and (-5, -6, 19) mm, and the coordinate of Model 4 was set as (-2, 5, 6) mm and (3, 6, 9) mm. The experimental results are shown in Fig 4, and the quantitative analysis is shown in Table III. The experimental results show that FISTA-NET has better positioning accuracy and shape recovery ability in dual-source environment.

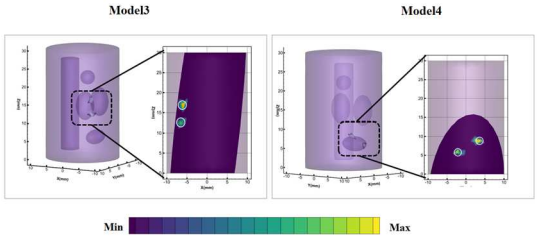

Fig 4. is the 3D view and 2D cross section of the reconstruction results of Model 3 and Model 4 using FISTA-NET method. As for the selection of 2D cross section, we need to ensure that can pass through the dual-source at the same time, and the white circle represents the position of the true Cerenkov source.

TABLE III. QUANTITATIVE RESULTS OF DUAL-SOURCE

source	Coordinate of Cerenkov source(mm)	LE	DICE
Model 3	(-6 -5 18), (-5 -6 19)	1.491	0.622
Model 4	(-2 5 6), (3 6 9)	1.338	0.705

C. *Comparison of reconstruction results between FISTA-Net and FISTA*

In this section, to verify whether FISTA-NET can perform better than traditional FISTA in CLT, we compare the experimental results of FISTA method with those of FISTA-NET. The experimental results of FISTA are shown in Fig 5, which illustrates the 3D view and 2D cross-section of the reconstructed results. Fig 6 shows the quantitative comparative results of the two methods, in which FISTA-NET has lower LE value and higher DICE value, indicating that FISTA-NET has better positioning accuracy and shape recovery ability. It also can be concluded that learning the key parameters in FISTA by neural network has better performance than artificial setting.

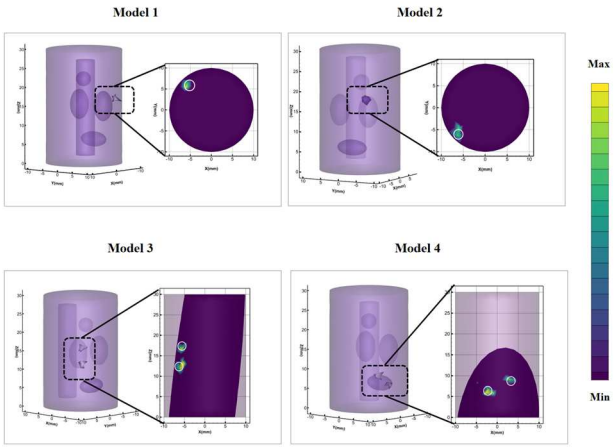

Fig 5. is the 3D view and 2D cross section of the reconstruction results of all Models using FISTA method.

V. CONCLUSION

In this paper, FISTA-NET is proposed to solve the inverse problem in CLT reconstruction. This method is a depth algorithm expansion based on FISTA and combines the advantages of model-based approach and neural network. Regularization parameters and threshold parameters of traditional FISTA need to be set manually, but FISTA-NET can dynamically learn related parameters. The previous neural network method is a black box method. Compared with the previous neural network method, this method is interpretable.

In order to verify the performance of FISTA-NET, numerical simulation experiments of single-source and dual-source were carried out. The experimental results show that FISTA-NET has good performance in positioning accuracy and shape recovery ability.

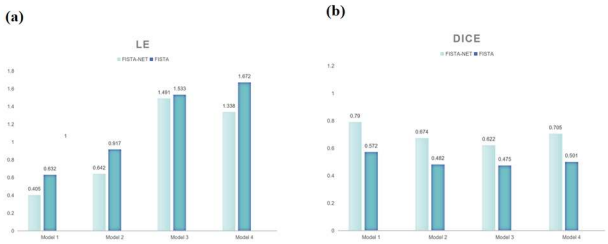

Fig 6. shows the quantitative comparison of FISTA-NET and FISTA reconstruction results.

However, FISTA-NET still has some shortcomings. At present, we have only conducted numerical simulation experiments without *in vivo* experiments to verify the feasibility of this method in living mice. In future work, we will conduct *in vivo* experiments to verify the feasibility of FISTA-NET *in vivo*.

REFERENCES

[1] P. A. Cherenkov, "Visible luminescence of pure liquids under the influence of γ-radiation," 1934.

[2] Changqing, Li, Gregory, S., Mitchell, Simon, R., and Cherry, "Cerenkov luminescence tomography for small-animal imaging," *Optics Letters,* vol. 35, no. 7, 2010.

[3] Z. Hu, J. Liang, W. Yang, W. Fan, C. Li, X. Ma, X. Chen, X. Ma, X. Li, and X. Qu, "Experimental Cerenkov luminescence tomography of the mouse model with SPECT imaging validation," *Optics Express,* vol. 18, no. 24, pp. 24441-24450, 2010.

[4] A. E. Spinelli, C. Kuo, B. W. Rice, R. Calandrino, and F. Boschi, "Multispectral Cerenkov luminescence tomography for small animal optical imaging," *Optics Express,* vol. 19, no. 13, pp. 12605-18, 2011.

[5] H. Liu, X. Yang, T. Song, C. Bao, L. Shi, Z. Hu, K. Wang, and J. Tian, "Multispectral hybrid Cerenkov luminescence tomography based on the finite element SPn method," 2015.

[6] Z. Zhang, M. Cai, Y. Gao, X. Shi, X. Zhang, Z. Hu, and J. Tian, "A novel Cerenkov luminescence tomography approach using multilayer fully connected neural network," *Physics in Medicine & Biology,* vol. 64, no. 24, pp. 245010 (10pp), 2019.

[7] X. Cao, X. Wei, F. Yan, L. Wang, and X. He, "A Novel Stacked Denoising Autoencoder-Based Reconstruction Framework for Cerenkov Luminescence Tomography," *IEEE Access,* vol. PP, no. 99, pp. 1-1, 2019.

[8] H. Meng, Y. Gao, X. Yang, K. Wang, and J. Tian, "K-nearest Neighbor Based Locally Connected Network for Fast Morphological Reconstruction in Fluorescence Molecular Tomography," *IEEE Transactions on Medical Imaging,* vol. PP, no. 99, pp. 1-1, 2020.

[9] X. Zhang, M. Cai, L. Guo, Z. Zhang, B. Shen, X. Zhang, Z. Hu, and J. Tian, "Attention mechanism-based locally connected network for accurate and stable reconstruction in Cerenkov luminescence tomography," *Biomedical optics express,* vol. 12 12, pp. 7703-7716, 2021.

[10] A. Beck, and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *Siam J Imaging Sciences,* vol. 2, no. 1, pp. 183-202, 2009.

[11] J. Xiang, Y. Dong, and Y. Yang, "FISTA-Net: Learning A Fast Iterative Shrinkage Thresholding Network for Inverse Problems in Imaging," 2020.