

基于 Unet 的切连科夫激发的荧光扫描断层重建

张文倩^{1,2}, 冯金超^{1,2*}, 李哲^{1,2}, 孙中华^{1,2}, 贾克斌^{1,2}¹北京工业大学信息学部计算智能与智能系统北京市重点实验室, 北京 100124;²先进信息网络北京实验室, 北京 100124

摘要 切连科夫激发的荧光扫描成像(CELSI)是一种新兴的成像技术,在生物医学领域有着广阔的应用前景。本课题组前期分别基于 Tikhonov 正则化和稀疏正则化实现了 CELSI 断层成像,但重建图像的质量仍有待提高。基于此,本文提出了一种基于 Unet 的图像后处理算法。该算法将 Tikhonov 方法只迭代一次得到的低质量图像作为 Unet 网络的输入,通过网络学习对重建的图像进行一定的修正,以进一步提高重建图像的质量。实验结果表明,所提算法可以提高计算效率,而且当荧光目标在深度 50 mm 处时重建图像的峰值信噪比和结构相似度分别能达到 28 dB 和 0.92。所提算法基于单荧光目标数据集训练的网络模型能够较好地重建出多荧光目标,表明了该算法具有较好的泛化能力。

关键词 医用光学;切连科夫激发的荧光扫描成像;断层成像;Tikhonov 正则化;Unet 神经网络

中图分类号 TP391

文献标志码 A

doi: 10.3788/CJL202148.1707001

1 引言

世界卫生组织/国际癌症研究署(WHO/IARC)发布的《2020 全球癌症报告》显示,2018 年全球新发癌症例数为 1810 万,死亡 960 万例^[1]。癌症作为全球第二大死亡原因,死亡例数和发病例数逐年上升,严重威胁着人类的生命健康。早期发现、早期诊断和精准治疗是提高恶性肿瘤治疗疗效的关键。分子影像作为一门新兴的前沿交叉学科,已成为早期肿瘤诊断、指导治疗和效果评估等的重要工具,在生物医学领域具有良好的发展前景^[2-4]。作为一种高分辨率在体成像技术,切连科夫激发的荧光扫描成像(CELSI)^[5]在生物医学领域吸引了研究人员的广泛关注。

2015 年,美国达特茅斯学院的 Brain W. Pogue 教授首次提出了 CELSI 的概念^[5]。CELSI 利用放疗直线加速器产生的 X 射线光片扫描生物体,由于切连科夫效应会产生切连科夫辐射,切连科夫辐射二次激发生物体内被荧光探针标记的分子或细胞并

产生比切连科夫光波长更长的荧光,荧光穿透生物组织传到生物体表面被外部的增强型 CCD 探测器接收^[6]。一方面,由于 CELSI 采用非常薄的 X 射线光片进行扫描,因此具有非常高的空间分辨率和深度灵敏度;另一方面,CELSI 采用切连科夫辐射作为内在激发光源,不仅具有较高的成像灵敏度,还能有效克服外在激发光源穿透能力差的问题^[7]。该技术是一种新兴的光学分子影像技术,具有空间分辨率高和成像深度大等独特优势。

但 CELSI 是一种二维成像技术,只能反映生物体表面的投影信息,无法获得荧光目标的三维空间分布情况。为了克服这一缺陷,本课题组在前期提出了切连科夫激发的荧光扫描断层成像(CELST)^[8]技术。

由于二次激发产生的荧光光子在到达生物体表面经历了多次散射,因此 CELST 重建的图像具有严重的病态性^[8]。同时,由于 CCD 探测器在采集荧光信号时会不可避免地混有噪声,因此 CELST 图像重建是一个典型的不稳定问题。

收稿日期: 2021-01-11; **修回日期:** 2021-02-19; **录用日期:** 2021-03-11

基金项目: 国家自然科学基金(81871394)、北京市教委面上项目(KM201810005030)

通信作者: *fengjc@bjut.edu.cn

在前期的工作中,本课题组基于 Tikhonov 正则化实现了 CELST 图像重建,但 Tikhonov 正则化重建的图像过于平滑^[9]。为进一步提高 CELST 重建图像的质量,本课题组将荧光目标的稀疏特性作为先验信息融入图像重建中,提出了基于近似消息传递(AMP)的稀疏正则化重建方法^[10]。相比 Tikhonov 正则化,基于 AMP 的方法可以减少重建图像的伪影,改善重建图像的质量^[10]。但当荧光目标位于生物体内较深位置时,该方法重建的图像仍会存在严重的伪影。因此,如何提高 CELST 重建图像的质量需要进一步研究。

近年来,随着深度学习技术的快速发展,基于深度学习的图像去噪和图像恢复方法得到了广泛关注。Jain 等^[11]采用卷积神经网络(CNN)对自然图像进行去噪,得到了比传统方法更优的去噪结果。Xu 等^[12]利用深度卷积神经网络进行图像去卷积实现了图像的复原,并获得了比传统方法更好的峰值信噪比(PSNR)和视觉效果。Shelhamer 等^[13]采用全卷积网络(FCN)对图像进行像素级分类,解决了语义级别的图像分割(semantic segmentation)问题。进一步,Ronneberger 等^[14]在 FCN 的基础上提

出了 Unet 网络,并将其用于医学图像的分割,提高了分割速度和分割准确率。Unet 网络结构简单,而且在图像处理方面具有一定优势,因此已被广泛应用于医学图像处理 and 图像重建中。

为了改善 CELST 的成像质量,本课题组将深度学习与传统的图像重建算法相结合,提出了一种基于 Unet 的图像后处理算法。该算法将基于 Tikhonov 正则化方法第一次迭代重建的低质量图像作为 Unet 网络的输入,通过网络训练去除重建图像中的伪影,提高了重建图像的质量。为验证本文所提算法的有效性,进行了数值仿真实验,并将所提算法与最新的基于近似消息传递(AMP)^[10]方法和基于图-总变分(Graph-TV)^[15]的正则化方法进行了比较。实验结果表明,本文算法不仅可以得到较为清晰准确的重建图像,而且具有计算效率高的优点。

2 基本原理

2.1 CELST 的数学原理(前向模型)

CELST 过程可以用耦合的扩散近似方程进行描述^[8,16],其数学表达式为

$$\begin{cases} -\nabla D_x(r) \nabla \Phi_x(r) + \mu_{ax}(r) \Phi_x(r) = S(r) \\ -\nabla D_m(r) \nabla \Phi_m(r) + \mu_{am}(r) \Phi_m(r) = \Phi_x(r) \eta \mu_{af}(r) \end{cases}, \quad (1)$$

式中:下标 x, m 分别表示直线加速器产生切连科夫辐射的过程和切连科夫辐射二次激发荧光目标产生荧光的过程; $\Phi(r)$ 是生物组织中 r 处的光流密度; $\mu_a(r)$ 是 r 处的光学吸收系数; $S(r)$ 是切连科夫辐射光能量密度; $D(r)$ 是 r 处的光学扩散系数, $D(r) = (\mu_a + \mu'_s)/3$,其中 $\mu'_s(r)$ 是 r 处的约化散射系数; $\mu_{af}(r)$ 是荧光目标对切连科夫辐射光的吸收系数; η 是荧光目标的量子效率; $\eta \mu_{af}(r)$ 是荧光量子产额。

采用有限元方法(FEM)^[16]对方程(1)进行离散求解,可以得到表面荧光信号和生物体内未知的荧光量子产额分布之间的线性关系为

$$\mathbf{A}x = \mathbf{y}_{\text{meas}}, \quad (2)$$

式中: $\mathbf{A}_{m \times n}$ ($m \ll n$) 是系统矩阵,可以通过求解耦合的扩散方程得到; \mathbf{y}_{meas} 是表面测得的荧光信号; x 是未知的荧光量子产额。

2.2 基于 Unet 的 CELST 图像重建算法

由于光在生物组织中经历了多次散射且采集的荧光信号 \mathbf{y}_{meas} 混有噪声,因此(2)式是一个严重的

病态问题,不能直接进行求解。为求解(2)式,本课题组基于 Tikhonov 正则化将 CELST 图像重建转化为求解(3)式所示的最优化问题。

$$x = \arg \min_x \frac{1}{2} \|\mathbf{A}x - \mathbf{y}_{\text{meas}}\|_2^2 + \lambda \|x\|_2^2, \quad (3)$$

式中: λ 为正则化参数,用来平衡数据拟合项 $\|\mathbf{A}x - \mathbf{y}_{\text{meas}}\|_2^2$ 和正则项 $\|x\|_2^2$ 。

Tikhonov 正则化重建的荧光目标过于光滑,而且随着荧光目标深度的增加,重建的图像中存在明显的伪影。为了解决正则化方法重建图像质量差的问题,本课题组基于 Unet 网络对 Tikhonov 正则化重建的图像进行后处理。为提高整个算法的计算速度,将 Tikhonov 正则化第一次迭代重建的低质量图像作为 Unet 网络的输入。图 1 展示了所提算法的框架。

如图 1 所示,所提算法主要包括两个阶段:

第一阶段:初步图像的重建。采用 Tikhonov 正则化方法迭代一次得到初步的荧光量子产额分布。

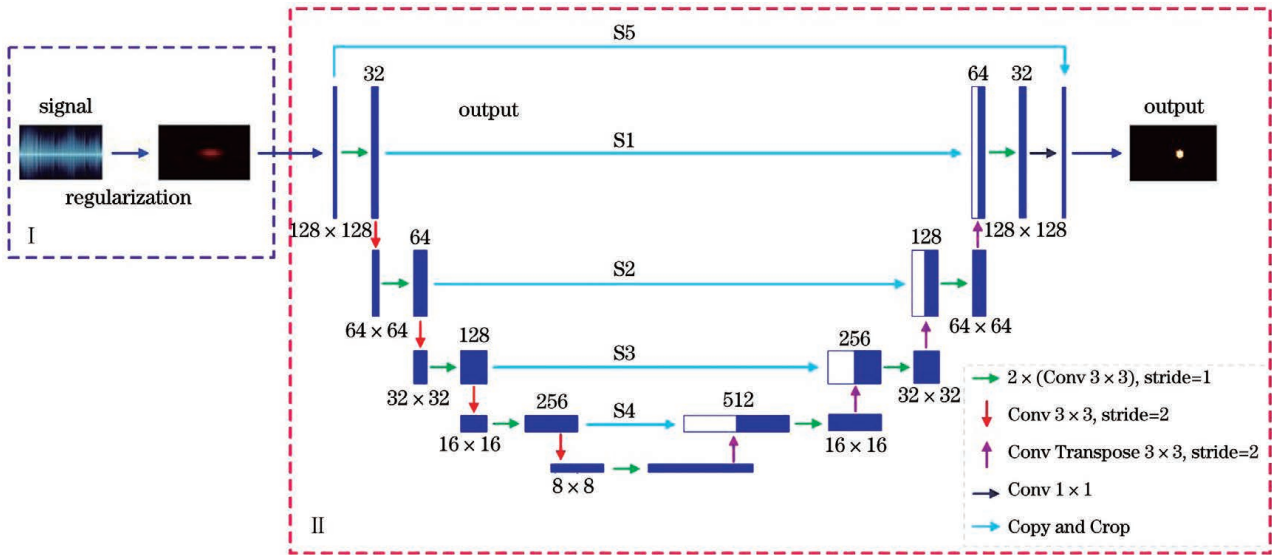


图 1 所提算法示意图

Fig. 1 A schematic representation of the proposed algorithm

第二阶段:图像后处理。将第一阶段重建出的低质量图像作为 Unet 网络的输入,通过网络的训练学习去除重建图像中的伪影,得到高质量图像。如图 1 所示,Unet 网络由编码器和解码器组成。编码器包含 4 个卷积块和 4 个下采样层,每个卷积块包含 2 个卷积核大小为 3×3 、步长为 1 的卷积层。对于下采样层,由于最大/最小池化 (Max/Min pooling) 或平均池化的滤波器系数在网络训练中是常量,可能会造成信息损失^[17],进而导致重建的图像不理想。为了提高重建图像的质量,本文采用卷积核大小为 3×3 、步长为 2 的卷积层实现下采样。解码器包含 4 个卷积块和 4 个上采样层,其中,卷积块同编码器的卷积块。对于网络的上采样,为避免插值带来的信息不准确的问题,本文采用卷积核大小为 3×3 、步长为 2 的转置卷积 (transposed convolution) 替换标准 Unet 中采用的直接插值。网络的卷积层均使用 leaky ReLU 作为激活函数,采用批量归一化 (BN) 加速学习。为提升网络的性能以及增加网络可学习的特征数量,本文在标准 Unet 网络的基础上,在第一层神经网络与最后一层神经网络之间增加了跳跃连接。

进一步,为了获得稳定清晰的重建结果,本文对 Unet 网络中采用的损失函数进行了改进。不同的损失函数会对图像质量造成不同的影响,人眼的主观感受也不一样^[18]。 L_2 损失函数不仅会放大最大误差和最小误差之间的差距,还会对异常点比较敏感;而 L_1 损失函数能较好地保留亮度和颜色信息,鲁棒性好,但不稳健^[19]。因此,在训练网络模型时,

仅用 L_1 损失函数或者 L_2 损失函数会很容易产生局部最优解。由于结构相似度 (SSIM) 损失函数综合考虑了亮度 (luminance)、对比度 (contrast) 和结构 (structure) 指标^[20],因此可以保留重建图像的高频信息 (图像的边和细节),但它容易导致亮度的改变和颜色的偏差。基于此,本文结合 L_1 损失函数和 SSIM 损失函数设计了一种混合损失函数,即

$$L_{\text{Mix}} = \alpha L_{\text{SSIM}} + (1 - \alpha) L_{l_1}, \quad (4)$$

式中: L_{SSIM} 和 L_{l_1} 分别是 SSIM 和 L_1 损失函数; α 是权重因子,其值由经验确定。在本文中, α 的值设为 0.84。

SSIM 损失函数 L_{SSIM} 的定义为^[20]

$$L_{\text{SSIM}} = 1 - \frac{(2\mu_x\mu_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2)}, \quad (5)$$

式中: x 是真实值; \hat{x} 是模型输出的重建值; μ_x 、 $\mu_{\hat{x}}$ 、 σ_x 、 $\sigma_{\hat{x}}$ 分别是 x 和 \hat{x} 的均值和标准差; $\sigma_{x\hat{x}}$ 表示 x 和 \hat{x} 的协方差。根据前期研究和多次实验,本文将 C_1 设定为 0.022, C_2 设定为 0.042。

L_1 损失函数的定义为

$$L_{l_1} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|, \quad (6)$$

式中: N 是图像中像素的个数。

2.3 数据集

本文采用大小为 $100 \text{ mm} \times 60 \text{ mm}$ 的仿体进行实验,它由 2747 个有限元节点和 5280 个三角形网格单元组成。根据 CELSI 的概念,在采集荧光信号时,将 CCD 探测器放置在相对于 X 射线 90° 的位

置^[5-6]。在本文的实验中,采用 X 射线光片等间距扫描 32 次,并在仿体上端放置了 64 个探测器,模拟

表面荧光信号的采集。仿体的背景的光学参数如图 2 所示^[10]。

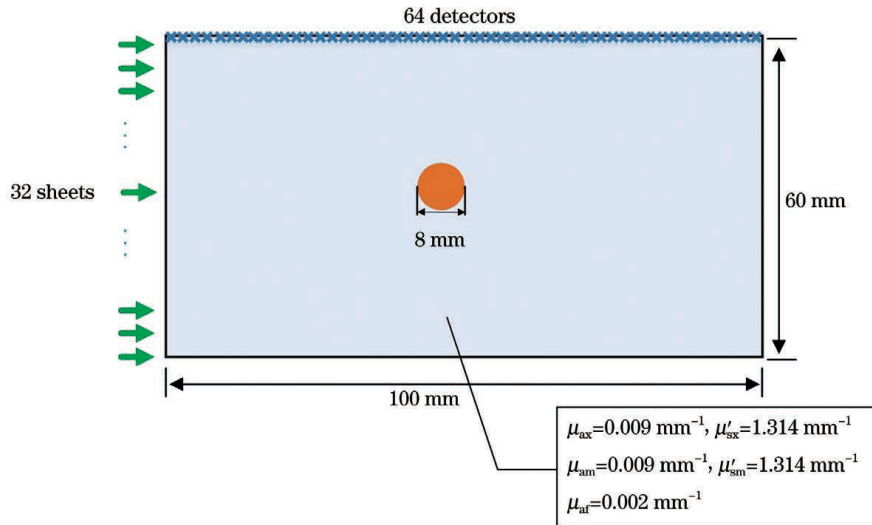


图 2 仿体及实验设置

Fig. 2 Phantom and experimental setting used in this study

为产生网络训练所用的数据集,通过改变荧光目标在仿体内的位置、荧光目标的个数、荧光目标与仿体背景间的荧光量子产额对比度得到不同的仿体,并通过求解(1)式得到三组实验数据集,将这三组数据集用于模型的训练、验证和测试。需要注意的是,测试集所用数据从未在网络训练中使用过。

第一组数据集中仅含有固定大小(半径为 4 mm)和固定对比度(4:1)的单个荧光目标,通过改变荧光目标在仿体内的位置得到 8364 个仿体以及相应的边界测量值。其中,6000 个样本数据用于训练,1200 个样本数据用于验证,1164 个样本数据用于测试。在 1164 个测试数据中,50 个不同深度的样本数据用于展示主观评价结果,其余 1114 个样本数据用于展示统计结果。

第二组数据集中含有两个半径为 4 mm、荧光

目标与背景荧光量子产额对比度为 4:1、边界距离不同的荧光仿体。通过改变两个荧光目标的位置和边界距离得到 8364 个数据对。同样,6000 个样本数据用于训练,1200 个样本数据用于验证,1164 个样本数据用于测试。在 1164 个测试数据中,挑选了深度 30 mm 处的 50 个数据用于展示主观评价结果,其余 1114 个数据用于展示统计结果。

第三组是混合数据集,通过改变仿体中荧光目标的荧光量子产额、位置和多目标间的距离,得到了 400 个数据对,其中包括不同对比度的单目标数据,不同距离、不同深度以及不同角度的双目标数据。第三组数据集只用作测试,以验证所提算法的泛化性。

图 3 展示了代表性的用于网络训练的实验数据,其中包括深度不同或荧光目标与背景荧光量子产额对比度不同的单荧光目标仿体、双荧光目标仿

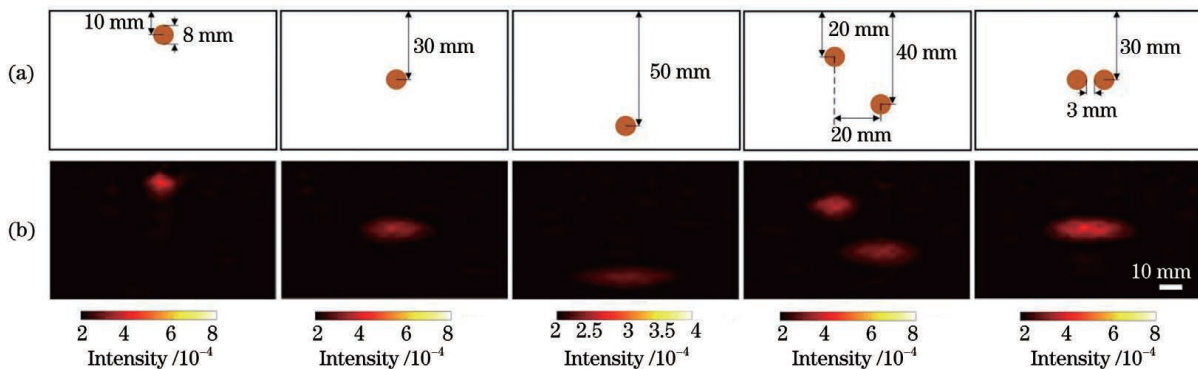


图 3 代表性的用于网络训练的实验数据。(a)代表性的真实图像;(b)代表性的 Unet 网络输入图像

Fig. 3 Representative experimental data used for network training. (a) Typical examples of true images; (b) typical examples of input images of the Unet network

体,以及边界距离不同的双荧光目标仿体。

为了展示所提算法 (Tikhonov-Unet) 的有效性,将其与 Tikhonov、AMP 和 Graph-TV 正则化算法进行了比较。为了量化实验结果,本文采用均方误差 (MSE)、峰值信噪比 (PSNR) 和结构相似性 (SSIM) 作为评价标准。

3 实验与分析

本文所用的计算机配置为 Inter(R) Core(TM) i7-9700 3.00 GHz CPU 和 2 个 GTX2080 GPU (8 GB 显存),操作系统为 Ubuntu,基于 Pytorch

1.2.0 搭建网络框架,在 CUDA 10.0 架构平台上进行计算。训练时采用 Adam 算法进行优化,通过多次实验确定初始学习率为 1×10^{-4} ,迭代次数为 100, Batch size 设为 16。经多次实验,将 Tikhonov 正则化参数设为 0.001。图 4 是训练过程中的损失函数曲线,其中实线代表训练集 (training dataset) 的损失函数,虚线代表验证集 (validation dataset) 的损失函数。从图 4 中可以看出,无论是单目标训练集还是多目标训练集,网络经过训练后都能快速收敛。另外, Tikhonov、AMP 和 Graph-TV 三种正则化算法中所用的正则化参数均通过多次实验后选择最优值。

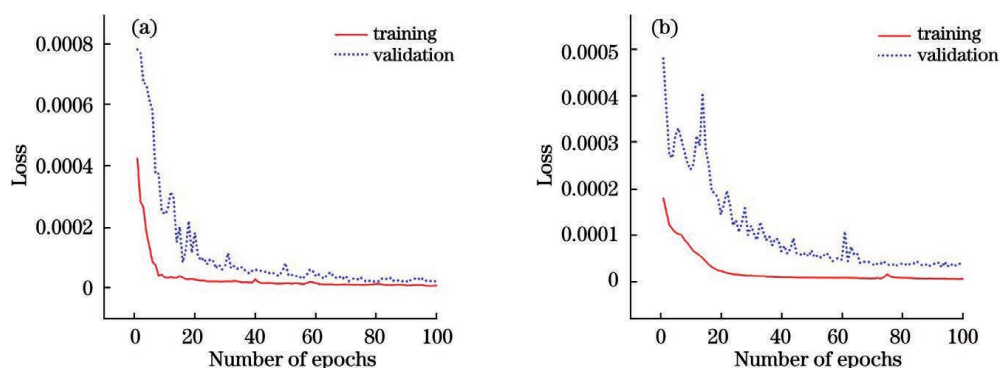


图 4 训练阶段训练集和验证集下的损失曲线。(a)单目标数据集下的损失曲线;(b)双目标数据集下的损失曲线

Fig. 4 Loss curves for training and validation datasets during the training stage. (a) Loss curves for single target dataset; (b) loss curves for two targets dataset

3.1 单荧光目标实验结果

首先,利用单荧光目标训练得到的网络模型来

重建单荧光目标图像。图 5 显示了荧光目标位于仿体不同深度(从 10 mm 到 50 mm)时的代表性结

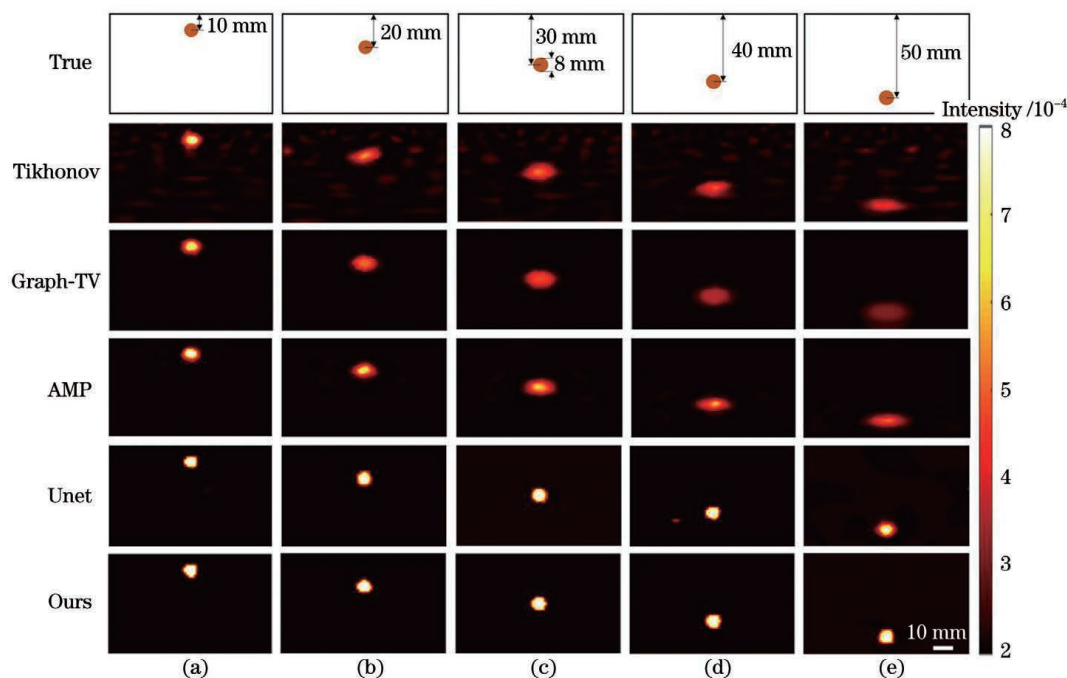


图 5 不同深度时不同算法的重建结果,(a)~(e)中荧光目标的深度从 10 mm 以步长 10 mm 均匀增加到 50 mm

Fig. 5 Reconstructed images by different algorithms at different depths, where the depth of the fluorescent target in (a)~(e) increases from 10 mm to 50 mm with the step size of 10 mm

果,从上到下依次为真值(True)图像以及 Tikhonov 正则化(迭代 5 次收敛)、Graph-TV 正则化、AMP、标准 Unet、本文算法(Ours)的重建结果。

从图 5 可以看出:当荧光目标中心距离表面深度为 10 mm 时,4 种算法都可以准确地重建荧光量子产额的分布,但 Tikhonov 重建的图像含有明显的伪影,因此 PSNR (23.34 dB)和 SSIM (0.53)最低;Graph-TV 和 AMP 重建的荧光目标发生了形变。图 6 显示了沿荧光目标中心位置水平方向的荧光产额值的轮廓线。综合图 5 和图 6 可以看出,相比 Tikhonov、Graph-TV 和 AMP,本文算法获得了最好的结果,重建出的荧光目标在量子产额、位置、大小上都最接近真实值。

从图 5 中还可以看到,随着荧光目标在仿体中深度的增加,Tikhonov、Graph-TV 和 AMP 三种算法的性能逐渐降低,而本文算法仍能保持较好的重建性能。相比 Tikhonov 算法和标准 Unet 的重建结果,本文算法采用改进 Unet 网络的重建结果可以获得更准确的荧光产额值,在深度 50 mm 处的重建结果优势明显。

图 7(a)~(c)展示了 4 种算法的量化结果。由图 7(a)可见,本文算法在 MSE 值上比其他方法要小一个数量级。就 PSNR 而言,当荧光目标深度为 10 mm 时,Tikhonov 重建的 PSNR 和 SSIM 最低;相比 Graph-TV,本文算法的 PSNR 提高了 28%,即使当荧光目标深度为 50 mm 时,本文算法重建的

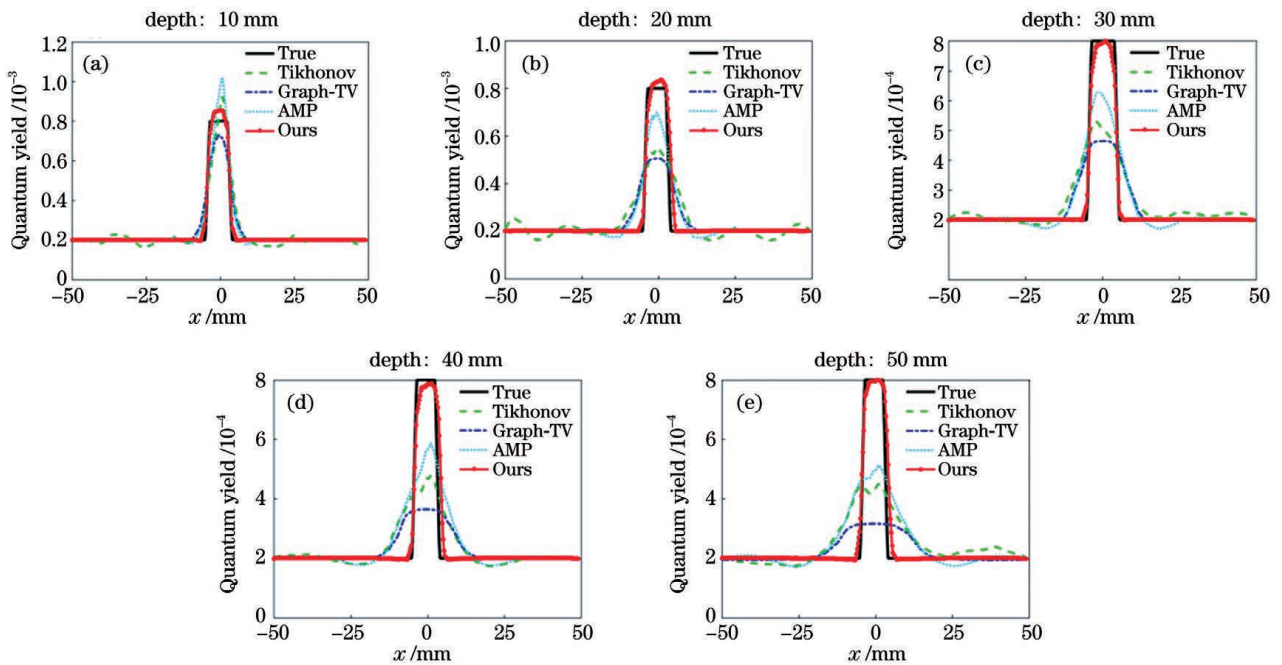


图 6 单目标重建图像沿荧光目标中心水平方向的荧光产额的轮廓线,(a)~(e)中荧光目标的深度从 10 mm 以步长 10 mm 均匀增加到 50 mm

Fig. 6 Profiles of quantum yield of single target reconstructed image along horizontal direction of the center of the fluorescence target, where the depth of the fluorescent target in (a)~(e) increases from 10 mm to 50 mm with the step size of 10 mm

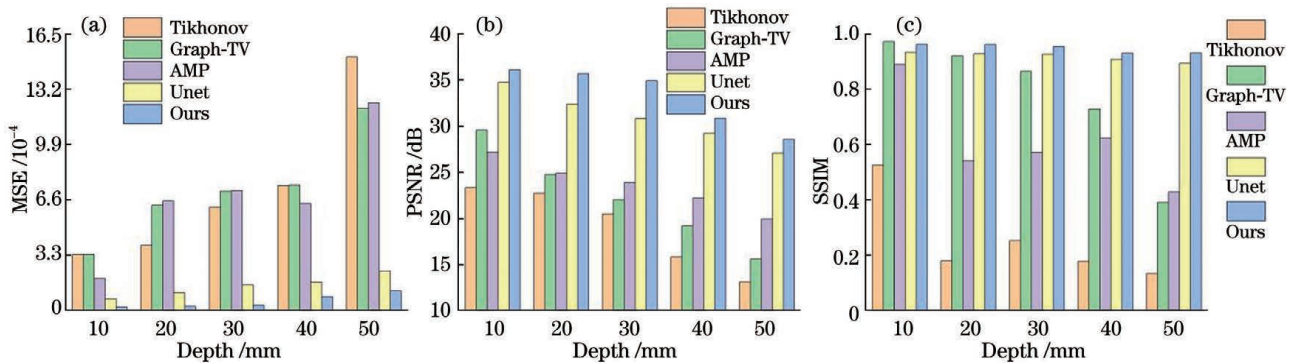


图 7 不同算法的量化结果。(a) MSE; (b) PSNR; (c) SSIM

Fig. 7 Quantitative results for different algorithms. (a) MSE; (b) PSNR; (c) SSIM

PSNR 值也能达到 28 dB,比 Tikhonov、Graph-TV 和 AMP 正则化算法分别提高了约 65%、56% 和 42%。同时,本文算法的 SSIM 值仍可达到 0.92,远远高于其他三种算法,进一步表明了本文算法的优势。

表 1 给出了各算法的重建时间。由于正则化算法都需要多次迭代,每次迭代都需要计算 Jacobian 矩阵(大小为 2048×2747)以及对 Hessian 矩阵(大

小为 2048×2048)求逆操作,因此计算时间较长;而本文算法只需要 Tikhonov 迭代重建一次,再代入网络模型前向计算一次,因此重建时间较短。三种正则化算法的重建时间都超过了 45 s,其中,相比另外两种正则化算法,Tikhonov 需要的计算时间较少,平均每次迭代需要 10 s 以上。本文算法平均需要 10.85 s,比 Tikhonov 至少提高了约 80%,大大提高了重建效率。

表 1 不同算法的重建时间比较

Table 1 Comparison of reconstruction time for different algorithms

Depth /mm	Reconstruction time /s				
	Tikhonov	Graph-TV	AMP	Unet	Ours
10	61.35	112.11	114.27	11.21	13.06
20	55.47	137.93	117.71	10.56	10.43
30	53.27	149.97	121.09	10.27	10.08
40	46.80	133.25	106.31	9.11	8.21
50	59.09	139.13	124.49	13.02	12.47

图 8 和表 2 展示了不同算法在整个测试数据集上的统计结果。该结果进一步表明本文算法相比其

他算法,可以更准确地重建不同深度处的荧光目标,具有较好的重建性能。

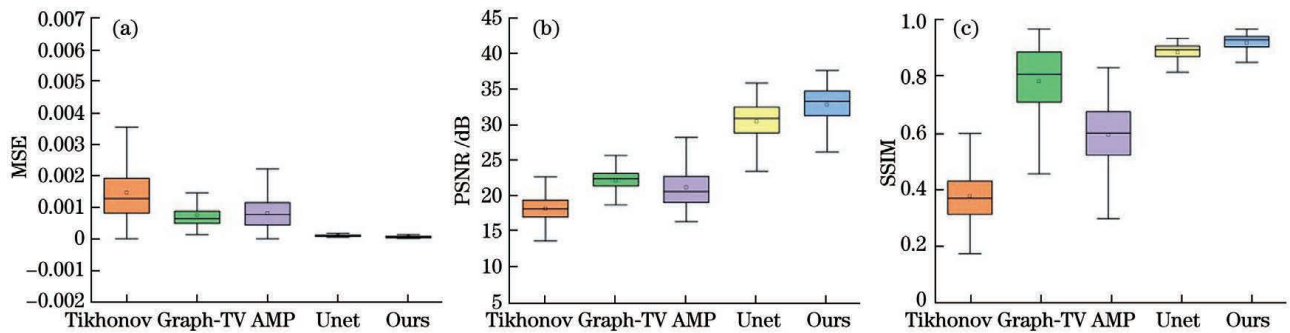


图 8 不同算法的统计结果。(a) MSE;(b) PSNR;(c) SSIM

Fig. 8 Statistic results for different algorithms. (a) MSE; (b) PSNR; (c) SSIM

表 2 不同算法的统计结果

Table 2 Statistic results for different algorithms

Algorithm	MSE	PSNR /dB	SSIM
Tikhonov	1.40×10^{-3}	18.135	0.383
Graph-TV	7.59×10^{-4}	22.095	0.794
AMP	8.18×10^{-4}	21.141	0.603
Unet	1.04×10^{-4}	30.422	0.897
Ours	6.15×10^{-5}	32.821	0.933

3.2 多荧光目标的实验结果

为了验证本文算法对多荧光目标的重建性能,本课题组在双荧光目标数据集上训练网络模型,然后进行多目标仿体的重建。图 9 展示了深度

30 mm 处两个半径为 4 mm 的荧光目标的重建结果。其中,两个荧光目标的边界距离从 5 mm 以步长 1 mm 逐渐减小到 0 mm。图 10 展示了深度为 30 mm 时沿荧光目标中心位置水平方向上的荧光产额的轮廓线。由图 9 和图 10 可以看出,当两个荧光目标的边界距离小于 5 mm 时,Tikhonov、Graph-TV 和 AMP 三种正则化算法不能区分两个荧光目标,而本文算法在边界距离为 1 mm 时仍可以分辨出两个荧光目标。

图 11 显示了基于训练的模型重建不同深度下的多个荧光目标仿体的结果。从图 11 可以看出,本文算法均可以较好地重建出不同深度下的 2 个、3 个或 4 个荧光目标,而且重建目标的位置、形状和

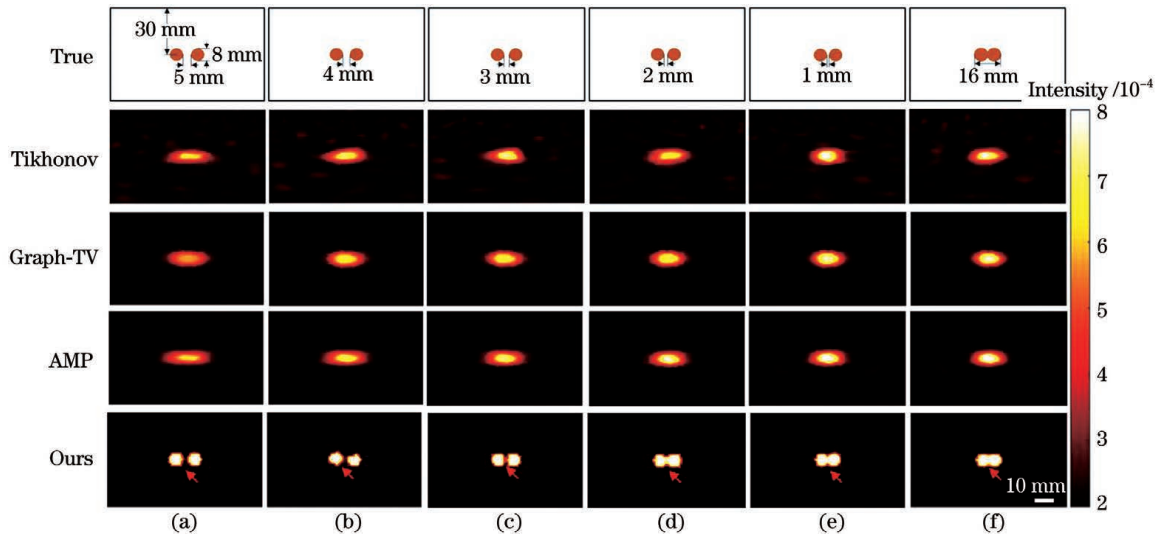


图 9 双荧光目标时不同算法的重建结果, (a)~(f) 中荧光目标的边界距离从 5 mm 以步长 1 mm 减小到 0 mm
 Fig. 9 Reconstructed images with different algorithms in case of two targets, where the edge-to-edge distance of two fluorescent targets in (a)–(f) decreases from 5 mm to 0 mm with step size of 1 mm

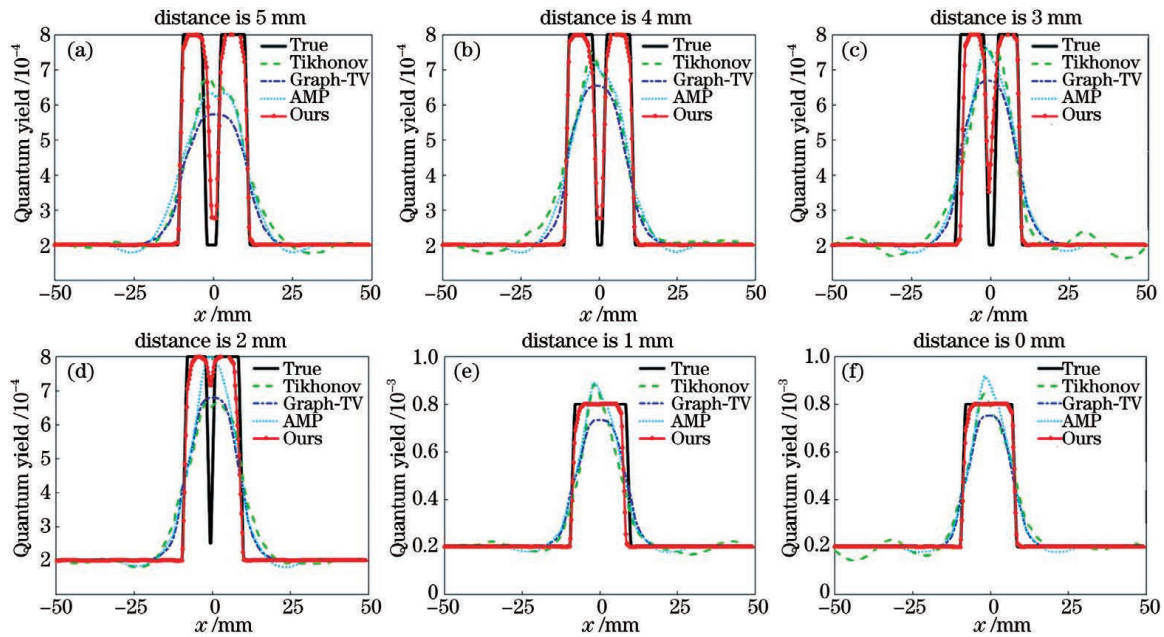


图 10 多目标重建图像沿荧光目标中心水平方向的荧光产额的轮廓线, (a)~(f) 中荧光目标的边界距离从 5 mm 以步长 1 mm 减小到 0 mm

Fig. 10 Profiles of quantum yield of multiple targets reconstructed image along horizontal direction of the center of the fluorescence target, where the edge-to-edge distance of two fluorescent targets in (a)–(f) decreases from 5 mm to 0 mm with the step size of 1 mm

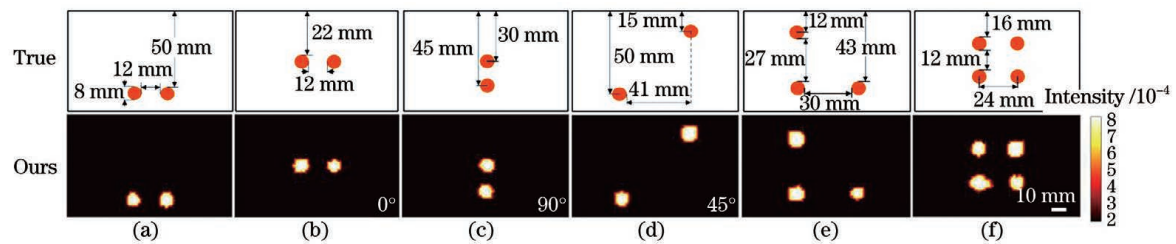


图 11 不同深度下本文算法的多目标重建结果

Fig. 11 Reconstructed images of multiple targets by our algorithm at varied depths

荧光产额值都接近真实值。这也表明了本文算法具有一定的泛化性。

3.3 泛化能力测试

为进一步验证 Unet 网络的泛化能力,本课题

组利用第一组实验中在单荧光目标数据集上得到的网络模型,重建不同对比度的单荧光目标图像和不同边界距离的多荧光目标图像,实验结果分别如图 12 和 13 所示。

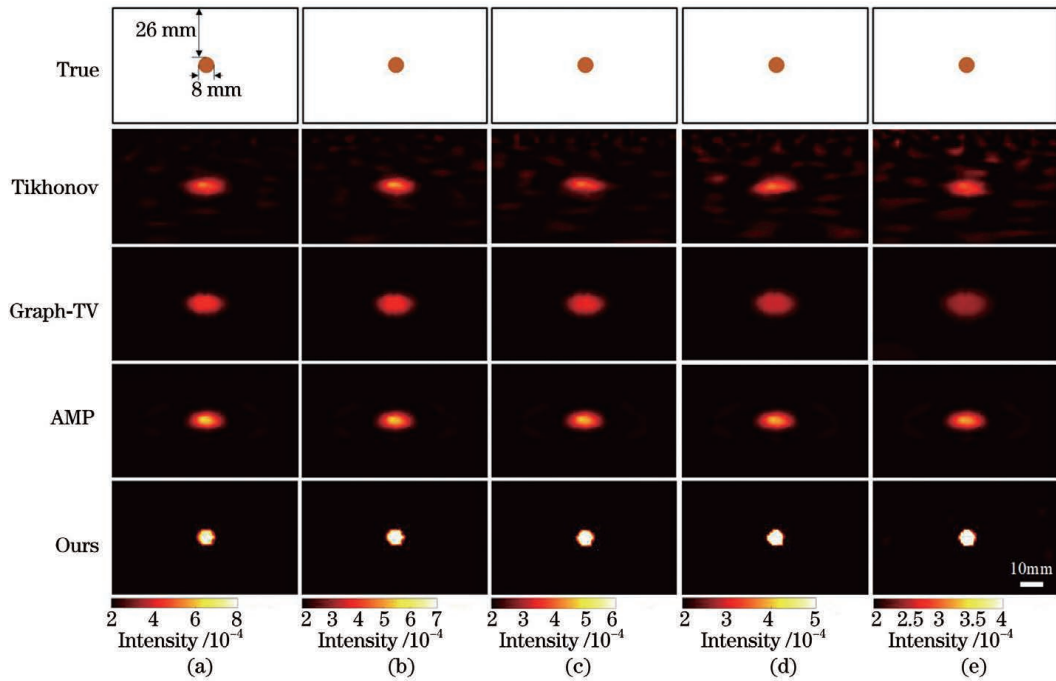


图 12 不同对比度下单荧光目标图像的重建结果。(a)对比度为 4:1;(b)对比度为 3.5:1;(c)对比度为 3:1;(d)对比度为 2.5:1;(e)对比度为 2:1

Fig. 12 Reconstructed images of single target at different contrasts. (a) Contrast is 4:1; (b) contrast is 3.5:1; (c) contrast is 3:1; (d) contrast is 2.5:1; (e) contrast is 2:1

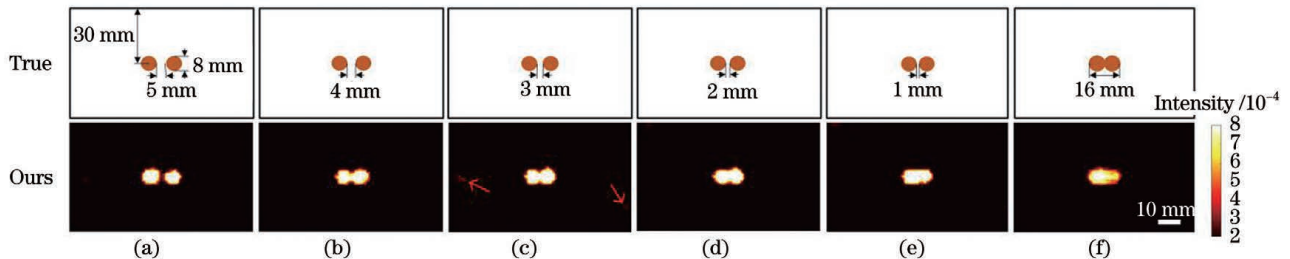


图 13 单荧光目标训练模型用于重建双荧光目标的结果。(a)边界距离为 5 mm;(b)边界距离为 4 mm;(c)边界距离为 3 mm;(d)边界距离为 2 mm;(e)边界距离为 1 mm;(f)边界距离为 0 mm

Fig. 13 Reconstructed images of two targets using the model trained on a single target datasets. (a) Edge-to-edge distance is 5 mm; (b) edge-to-edge distance is 4 mm; (c) edge-to-edge distance is 3 mm; (d) edge-to-edge distance is 2 mm; (e) edge-to-edge distance is 1 mm; (f) edge-to-edge distance is 0 mm

由图 12 可以看出,随着对比度的逐渐减小,基于正则化算法重建的图像质量逐渐降低。在荧光目标与背景的荧光量子产额对比度为 3:1 时, Tikhonov 算法重建的图像已经出现了较为明显的伪影, Graph-TV 和 AMP 两种算法不能准确重建荧光目标的形状和荧光产额值,而本文算法可以准确地重建荧光目标。当对比度为 3.5:1 时,相比 Graph-TV 和 AMP 算法,本文算法将 PSNR 和

SSIM 值至少提高了 38% 和 12%。当对比度为 4:1 时,相比 Tikhonov、Graph-TV 和 AMP,本文算法将 PSNR 和 SSIM 值分别提高到了 32.19 dB 和 0.93。上述结果表明,本文算法的 PSNR 和 SSIM 比 Tikhonov、Graph-TV 和 AMP 方法的平均值分别提高了 65% 和 91%。

图 13 显示了单荧光目标数据集训练的网络模型用于含两个荧光目标仿体的重建结果。由图 13

可以看出:当荧光目标的边界距离大于 3 mm 时,Unet 网络可以较好地分辨出两个荧光目标,但当荧光目标的边界距离小于 3 mm 时,Unet 网络就不能分辨出两个荧光目标。实验结果表明,用单目标数据集训练的 Unet 网络在分辨能力上相比多目标训练的模型有所衰减,但可以适用于边界距离约为 3 mm 的双荧光目标的重建。

4 结 论

为提高 CELST 的成像速度和重建图像的质量,本课题组针对前期图像重建算法中存在的问题,提出了基于 Unet 的 CELST 图像重建方法。该方法先基于 Tikhonov 方法迭代一次进行初步重建,将得到的低质量图像作为 Unet 网络的输入,经过网络训练后得到高质量的输出图像。

实验结果表明,与其他算法相比,本文算法提高了重建速度和重建图像的质量,不但可以分辨近距离的多个荧光目标,还可以使 30 mm 深度处的分辨率达到 1 mm。进一步的实验表明,本文方法在单目标训练集上训练得到的网络模型,在低对比度时可以较好地重建出多荧光目标,具有一定的泛化能力。

另一方面,本文算法在单目标上训练出的模型用于多目标重建时分辨能力有所减弱。拟在后续工作中对网络模块进行改进,以获得更好的重建质量和更高的泛化能力。此外,本文算法采用后处理方法进行 CELST 成像,没有直接对信号进行重建,后续将进行端到端实验,以避免数学建模引起的误差。

参 考 文 献

- [1] World Health Organization. World cancer report: cancer research for cancer prevention [R]. Lyon: WHO, 2020: 253.
- [2] Blankenberg F G. Molecular imaging with single photon emission computed tomography [J]. IEEE Engineering in Medicine and Biology Magazine, 2004, 23(4): 51-57.
- [3] Huang W, Li X, Morris E A, et al. The magnetic resonance shutter speed discriminates vascular properties of malignant and benign breast tumors *in vivo* [J]. Proceedings of the National Academy of Sciences of the United States of America, 2008, 105(46): 17943-17948.
- [4] Nienhaus G U, Wiedenmann J. Structure, dynamics and optical properties of fluorescent proteins: perspectives for marker development[J]. Chemphyschem, 2009, 10(9/10): 1369-1379.
- [5] Zhang R, D'souza A V, Gunn J R, et al. Cherenkov-excited luminescence scanned imaging [J]. Optics Letters, 2015, 40(5): 827-830.
- [6] Pogue B W, Feng J C, LaRochelle E P, et al. Maps of *in vivo* oxygen pressure with submillimeter resolution and nanomolar sensitivity enabled by Cherenkov-excited luminescence scanned imaging[J]. Nature Biomedical Engineering, 2018, 2(4): 254-264.
- [7] Xie L N, Gao C Y, Wang Q, et al. Research progress of photodynamic therapy based on Cherenkov radiation for tumors [J]. Laser & Optoelectronics Progress, 2020, 57(19): 190002. 谢丽娜, 高诚宜, 汪琪, 等. 基于切伦科夫辐射的光动力疗法用于肿瘤治疗的研究进展 [J]. 激光与光电子学进展, 2020, 57(19): 190002.
- [8] Feng J C, Bruza P, Dehghani H, et al. Cherenkov-excited luminescence sheet imaging (CELSI) tomographic reconstruction[J]. Proceedings of SPIE, 2017, 10049: 1004912.
- [9] Cao X, Zhang B, Wang X, et al. An adaptive Tikhonov regularization method for fluorescence molecular tomography [J]. Medical & Biological Engineering & Computing, 2013, 51(8): 849-858.
- [10] Feng J C, Chang D, Li Z, et al. Cherenkov-excited luminescence scanned tomography reconstruction based on approximate message passing [J]. Chinese Journal of Lasers, 2020, 47(2): 0207027. 冯金超, 常迪, 李哲, 等. 基于近似消息传递的切伦科夫激发的荧光扫描断层重建 [J]. 中国激光, 2020, 47(2): 0207027.
- [11] Jain V, Seung H S. Natural image denoising with convolutional networks [C] // International Conference on Neural Information Processing Systems, December 8-11, 2008, Vancouver, British Columbia, Canada. New York: Curran Associates Inc, 2008: 769-776.
- [12] Xu L, Ren J S J, Liu C, et al. Deep convolutional neural network for image deconvolution [C] // 28th Annual Conference on Neural Information Processing Systems 2014, NIPS 2014, December 8-13, 2014, Montreal, Quebec, Canada. Taipei: Neural Information Processing Systems Foundation, 2014: 1790-1798.
- [13] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [14] Ronneberger O, Fischer P, Brox T. U-Net:

- convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [15] Lu W Q, Duan J M, Orive-Miguel D, et al. Graph and finite element-based total variation models for the inverse problem in diffuse optical tomography[J]. Biomedical Optics Express, 2019, 10(6): 2684-2707.
- [16] Dehghani H, Eames M E, Yalavarthy P K, et al. Near infrared optical tomography using NIRFAST: algorithm for numerical model and image reconstruction[J]. Communications in Numerical Methods in Engineering, 2008, 25(6): 711-732.
- [17] Fu J, Dong J B, Zhao F. A deep learning reconstruction framework for differential phase-contrast computed tomography with incomplete data[J]. IEEE Transactions on Image Processing, 2020, 29: 2190-2202.
- [18] Liu D, Li Y C. Quality assessment of remote sensing images based on deep learning and human visual system[J]. Laser & Optoelectronics Progress, 2019, 56(6): 061101.
刘迪, 李迎春. 基于深度学习和人眼视觉系统的遥感图像质量评价[J]. 激光与光电子学进展, 2019, 56(6): 061101.
- [19] Zhao H, Gallo O, Frosio I, et al. Loss functions for image restoration with neural networks[J]. IEEE Transactions on Computational Imaging, 2017, 3(1): 47-57.
- [20] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.

Reconstruction for Cherenkov-Excited Luminescence Scanned Tomography Based on Unet Network

Zhang Wenqian^{1,2}, Feng Jinchao^{1,2*}, Li Zhe^{1,2}, Sun Zhonghua^{1,2}, Jia Kebin^{1,2}

¹ Beijing Key Laboratory of Computational Intelligence and Intelligent System, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;

² Beijing Laboratory of Advanced Information Networks, Beijing 100124, China

Abstract

Objective As a new molecular imaging modality, Cherenkov-excited luminescence scanned imaging (CELSI) has demonstrated a great potential, especially in radiation therapy diagnostics. However, it can not provide in-depth information about molecular probes. Therefore, it is necessary to develop tomographic algorithms for CELSI. However, reconstructing spatial distributions of luminescent sources from boundary measurements is a typical ill-posed problem. Our previous work has demonstrated the feasibility and effectiveness of using Tikhonov and sparse regularizations for CELSI reconstruction. However, the quality of reconstructed images will degrade when the luminescent source is located at deep positions. The objective of this study is to develop a reconstruction algorithm for CELSI to improve the quality of reconstructed images.

Methods A two-stage reconstruction algorithm is developed in this study. First, low-quality images were reconstructed on the basis of Tikhonov regularization at the first iteration. Then, the resultant images were input into a revised Unet network, which had an encoder-decoder architecture. The encoder comprised four convolution blocks and four downsampling layers. Every convolution block comprised two convolution layers, and each convolution had a kernel size of 3×3 with a stride of 1. The downsampling layer had a convolution of size 3×3 with a stride of 2. In upsampling networks, the transposed convolution with a kernel of 3×3 and stride of 2 was used to replace direct interpolation used in the standard Unet. Here, we used leaky rectified linear units as the activation function to intensify the network in each convolutional layer. The batch normalization technique was used to accelerate learning. In addition, skip connection was applied to connect the first and last layers. The feasibility of the algorithm was evaluated through numerical simulations. Training and test datasets were generated using an open-source software, NIRFAST. For comparison, our algorithm was compared with Tikhonov regularization, approximate message passing (AMP), and graph-total variation (Graph-TV).

Results and Discussions First, a single circular target with 8 mm diameter was placed within the phantom with varying depths ranging from 10 mm to 50 mm. Although the four algorithms could reconstruct the distribution of targets, severe artifacts were found in images reconstructed by Tikhonov regularization, and the shapes of reconstructed targets were changed for AMP and Graph-TV (Fig. 5). Additionally, the quality of reconstructed images by Tikhonov regularization, AMP, and Graph-TV degraded as the depth increased. By contrast, our results reveal that the quantitative accuracy of recovered distributions of luminescence sources could be significantly improved by the proposed algorithm, which achieved the best image quality with a high peak signal-to-noise ratio (>28 dB) and structural similarity (>0.92). Furthermore, experiments with two luminescent sources were conducted to evaluate the algorithm's performance. When the edge-to-edge distance of two luminescent sources was <5 mm, Tikhonov regularization, AMP, and Graph-TV failed to reconstruct the source distributions, while the proposed algorithm could distinguish the two sources even when the edge-to-edge distance was ~ 1 mm (Fig. 9). In addition, the size and fluorescence yield of the sources reconstructed using the proposed algorithms were very close to their real values (Fig. 11). Further, the generalizability of the proposed algorithm was evaluated using a network trained on single target datasets. Our results demonstrate that the proposed algorithm could reconstruct luminescent sources accurately even when the contrast between the source and background was reduced to 2:1 (Fig. 12). The two luminescent sources could be distinguished well when the edge-to-edge distance was >3 mm (Fig. 13). The computational efficiency for the four algorithms is also demonstrated. The three traditional reconstruction algorithms require >45 s, whereas our algorithm requires ~ 11 s (Table 1).

Conclusions A tomographic reconstruction algorithm for CELSI is proposed to reconstruct distributions of luminescence sources based on a trained Unet neural network. Numerical simulations are used to evaluate the performance of our proposed algorithm. Our results reveal that both the image quality and quantitative accuracy of reconstructed fluorescence yield can be improved using the proposed algorithm compared with the conventional Tikhonov, sparse, and total variation regularizations. The proposed algorithm can reconstruct the distributions of luminescent sources when the depth is <50 mm. Moreover, it can recover the size and fluorescence yield of luminescent sources with an edge-to-edge distance of 1 mm. Numerical simulations of multiple luminescent sources show that the proposed algorithm has good generalizability when trained on a dataset with a single luminescent source. Furthermore, it is computationally efficient. Although the network input of our algorithm is a low-quality image reconstructed using Tikhonov regularization, the images reconstructed using other algorithms can be taken as network input without any processing.

Key words medical optics; Cherenkov-excited luminescence scanned imaging; tomography; Tikhonov regularization; Unet neural network

OCIS codes 170.3010; 170.6960