

基于改进的 U-Net 神经网络的稀疏视角光声图像质量增强方法

王通¹, 董文德², 沈康^{3,4}, 刘松德^{3,4}, 刘文¹, 田超^{3,4*}

¹中国科学技术大学物理学院, 安徽 合肥 230026;

²南京航空航天大学自动化学院, 江苏 南京 211106;

³中国科学技术大学工程学院, 安徽 合肥 230026;

⁴精密科学仪器安徽普通高校重点实验室, 安徽 合肥 230026

摘要 在光声断层成像中, 通常利用超声换能器阵列接收光声信号, 其制造成本较高, 并且阵元数量对最终成像质量有重要影响。为了提升稀疏视角下光声重建的图像质量, 提出了一种基于改进的 U-Net 神经网络结构的稀疏视角光声图像质量增强方法, 该方法采用的改进的 U-Net 网络的特点在于通过添加连续卷积层替换跳接层, 提升编码器和解码器拼接特征的匹配度; 同时利用了基于多尺度结构相似性指数的损失函数对网络进行训练。基于仿体数据集和活体数据集的实验结果表明, 改进的 U-Net 网络具有很好的图像细节重建能力, 其所得的重建图像质量优于经典的 U-Net 网络。

关键词 医用光学与生物光学; 光声断层成像; 图像重建; 稀疏视角; 深度学习; U-Net

中图分类号 R318.51

文献标志码 A

doi: 10.3788/LOP202259.0617022

Sparse-View Photoacoustic Image Quality Enhancement Based on a Modified U-Net

Wang Tong¹, Dong Wende², Shen Kang^{3,4}, Liu Songde^{3,4}, Liu Wen¹, Tian Chao^{3,4*}

¹School of Physical Science, University of Science and Technology of China, Hefei, Anhui 230026, China;

²College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 211106, China;

³School of Engineering Science, University of Science and Technology of China, Hefei, Anhui 230026, China;

⁴Key Laboratory of Precision Scientific Instrumentation of Anhui Higher Education Institutes, Hefei, Anhui 230026, China

Abstract In photoacoustic tomography, an ultrasonic transducer array is usually used to receive photoacoustic signals, which is expensive to manufacture, and the number of array elements has an important impact on the final imaging quality. To improve photoacoustic image quality reconstructed under sparse view condition, this study proposes a modified U-Net based on the replacement of the skip connection in a conventional U-Net with continuous convolutional layers, thereby increasing the matching degree of features transferred from the encoder to the decoder. Furthermore, the loss function based on the structural similarity index measure is used to train the network.

收稿日期: 2021-11-29; 修回日期: 2021-12-30; 录用日期: 2022-01-17

基金项目: 国家自然科学基金(62122072, 12174368, 61705216, 61905112)、安徽省科技重大专项(18030801138)、之江实验室(2019MC0AB01)、中国科学院人才项目、统筹推进世界一流大学和一流学科建设专项资金

通信作者: *ctian@ustc.edu.cn

Experimental results based on simulation and *in vivo* dataset show that compared with the conventional U-Net, the modified U-Net achieves more image details and the quality of the reconstructed image is significantly better.

Key words medical optics and biotechnology; photoacoustic tomography; image reconstruction; sparse view; deep learning; U-Net

1 引言

光声断层成像是一种基于光声效应的非侵入式生物成像技术,其融合了光学成像的高对比度和超声成像在较深层组织中的高分辨率优势^[1],在早期癌症检测^[2]、细胞成像^[3-4]、活检引导^[5-6]等医学成像领域取得了日益广泛的应用。在光声成像过程中,脉冲激光照射生物组织,生物组织受热膨胀发出超声波,利用超声换能器阵列接收超声信号,再通过成像算法重建得到光声图像。为了得到高质量的图像,通常需要使用较多的换能器数量进行测量,但是更多的换能器数量会带来更高的制作成本、硬件复杂度以及对设备计算能力的更大挑战,因此需要对换能器数量进行限制,在稀疏视角情况下进行测量。稀疏视角光声成像可以有效降低系统复杂度,减少成本,同时减少需要处理的数据量加快成像速度,但是也会导致重建的光声图像质量下降,出现伪影和模糊现象^[7-8],在图像重建时需要考虑稀疏视角的影响。

传统的光声成像算法分为两类:1)基于光声成像原理直接重建的滤波反投影(FBP)算法^[9],该类算法采用逆Radon变换的方法直接求解光声方程得到初始声压分布,在测得的数据量足够多并且噪声很小的情况下可以重建得到精确的光声图像,在稀疏视角重建时有明显的伪影;2)基于模型的迭代求解算法,该类算法寻找最优解的方式是最小化测得的信号与通过前向模型预测得到的信号之间的误差,并且引入先验知识来对重建图像进行约束来提升图像质量,通过施加正则项的方法提高非理想条件下光声重建的图像质量^[10]。研究人员提出了很多不同的正则项形式来对图像进行约束,包括Tikhonov正则^[11-12]、全变分正则^[13-14],以及结合了不同类型的正则项形式的自定义正则^[15]等,有效提升了迭代重建的图像质量。但是,一方面正则项对复杂结构的重建质量提升有限,并且对应的正则项系数需要根据经验来进行人为选择,对最终效果有很大的影响;另一方面,形式复杂的正则项也会限制迭代所用优化算法的选择空间,并且增加运算量,所以这类方法也存在着一定的局限性。

近些年,卷积神经网络(CNN)在医学成像领域得到了广泛的关注,在生物医学图像分类^[16-17]、图像分割^[18-20]以及图像去噪^[21]等领域发挥了强大的作用,将CNN用于光声成像领域的研究也取得了令人瞩目的成果。Awasthi等^[22]采用一种改进的U-Net网络对带宽受限的光声信号进行处理,实现超分辨成像;Zhang等^[23]将AlexNet和GoogLeNet用于光声图像中进行图像分类和分割以实现癌症诊断;Tong等^[24]提出一种特征投影网络(FPNet),从稀疏采样和有限视角数据直接重建得到光声图像。其中U-Net是一种应用非常广泛的网络结构^[25],U-Net网络可以有效提取图像特征并同时去除重建图像中的伪影,在光声成像中已被用于数据预处理^[22],稀疏视角和有限视角等非理想条件下重建图像的后处理^[26-27],以及从数据直接重建图像^[28]等任务。但是,U-Net网络也存在一定的不足,主要表现在提取图像特征的过程中会存在细节丢失的问题。有学者针对这一经典结构提出了许多改进方案,比如引入残差网络结构^[29-30],用全连接层替代简单的最大池化层以提取更多的图像特征信息^[31-32],或者利用离散小波变换替代最大池化层以获取图像高频信息^[33-34]等。本文在保留U-Net主要架构的基础上对其进行了改进:一方面利用连续的卷积层替代经典U-Net中的跳接层;另一方面基于多尺度结构相似性指数设计了损失函数并对网络进行训练。本文将上述改进的U-Net用于提升稀疏视角光声重建的图像质量,经过仿真和实验数据集的验证,改进的U-Net网络较经典的U-Net网络具有更好的图像细节重建能力。

2 基本原理及网络模型

2.1 光声成像及迭代图像重建原理

在光声成像过程中,脉冲激光照射生物组织,生物组织受热膨胀发出超声波。声波的产生和传播过程可通过下面的波动方程描述,

$$\nabla^2 p(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2} = -\frac{\beta}{C_p} \frac{\partial}{\partial t} H(\mathbf{r}, t), \quad (1)$$

式中: $p(\mathbf{r}, t)$ 为位置 \mathbf{r} 处在时间为 t 时的光声信号; $H(\mathbf{r}, t)$ 为加热函数,表示生物组织在单位时间单位

体积内吸收的能量; c 、 C_p 、 β 分别为组织声速、定压比热容和热膨胀系数。本文采用迭代算法求解上式, 将其写成矩阵形式,

$$\mathbf{Ax} = \mathbf{b}, \quad (2)$$

式中: \mathbf{x} 为初始声压分布; \mathbf{A} 为系统矩阵; \mathbf{b} 为测得的光声信号。光声重建的目标是由测得的光声信号 \mathbf{b} 求解声压分布 \mathbf{x} , 求解目标可以写成如下形式的最优化问题:

$$\mathbf{x} = \arg \min_{\mathbf{x}} \|\mathbf{Ax} - \mathbf{b}\|_2^2. \quad (3)$$

本文采用共轭梯度最小二乘 (CGLS) 方法对 (3) 式进行求解^[35], CGLS 迭代过程中由第 k 步到第 $k+1$ 步的更新策略如下,

$$\mathbf{u}_k = \mathbf{r}_k + \frac{\|\mathbf{r}_k\|_2^2}{\|\mathbf{r}_{k-1}\|_2^2} \mathbf{u}_{k-1}, \quad (4)$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{\|\mathbf{r}_k\|_2^2}{\|\mathbf{A}\mathbf{u}_k\|_2^2} \mathbf{u}_{k+1}, \quad (5)$$

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \frac{\|\mathbf{r}_k\|_2^2}{\|\mathbf{A}\mathbf{u}_k\|_2^2} \mathbf{A}^T \mathbf{A}\mathbf{u}_k, \quad (6)$$

式中: \mathbf{u}_k 为梯度下降时第 k 步的搜索方向; \mathbf{r}_k 为第 k 次迭代后的残差。变量的初值为

$$k=0, x_0=0, \mathbf{u}_{k-1}=0, \mathbf{r}_0 = \mathbf{A}^T(\mathbf{b} - \mathbf{A}\mathbf{x}_0). \quad (7)$$

实际应用过程中会使用正则项, 引入图像本身的信息对求解过程进行约束, 本文采用 Tikhonov 正则^[12]。

2.2 改进的 U-Net 神经网络结构

在换能器数量充足的情况下, 可以重建得到较为理想的光声图像; 在稀疏视角情况下, 换能器数量不足, 测量得到的有效信息有限, 会导致重建出的光声图像质量下降, 出现条纹状伪影^[8]。已有研究表明可以使用卷积神经网络提升稀疏视角下的重建质量, 将非稀疏视角情况下重建得到的图像作为参考图像 \mathbf{y}^* , 稀疏视角下重建得到的退化图像作为初始图像 \mathbf{y}_0 , 神经网络的目标是建立由 x 到 y 的映射,

$$\mathbf{y}^* = F(\mathbf{y}_0; \Theta), \quad (8)$$

式中: Θ 为网络中的参数。

U-Net 网络最初用于图像分割, 后来被引入医疗图像重建领域用于图像降噪、去伪影等任务, 图 1(a) 为一个典型 U-Net 网络的组成结构。U-Net 网络整体呈现出对称性, 分为编码器和解码器两部分, 编码器的功能是从图像中提取出空间特征, 解码器的功能是根据编码器提取到的特征去重新构建出目标图像的细节。

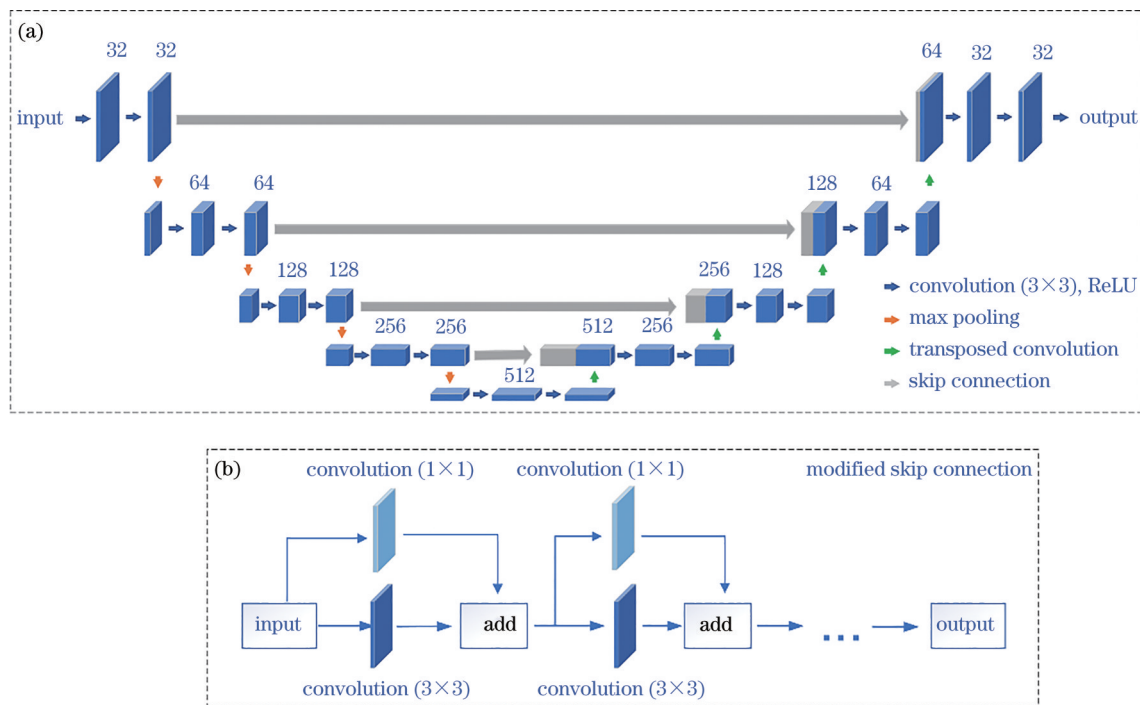


图 1 改进的 U-Net 网络结构。(a) 经典 U-Net 结构; (b) 改进的跳接层结构

Fig. 1 Architecture of the modified U-Net. (a) Conventional U-Net network; (b) modified skip connection network

U-Net最重要的设计之一在于跳接层的引入。U-Net中的编码器在编码过程中使用最大池化进行下采样操作以增加感受野并提取信息,同时减小图像尺寸以降低运算量,但是这一操作也会导致信息丢失。为了弥补信息丢失,跳接层将编码器每一层池化之前的结果拼接到解码器对应层中进行下一步操作,以尽量保留编码器中细节较为丰富的空间特征。但是,编码器中的特征与待拼接的解码器中的对应层并非完全匹配,这会限制整个网络的细节重建能力。以第一个跳接层为例,它将第一次池化操作之前的特征与最后一次反卷积操作之后的特征进行拼接,前者位于编码器的靠前的网络层,经过了较少的处理,提取的是较为低层次的特征;而后者位于解码器的靠后的网络层,经历了多层网络的处理,提取到的是高层次的特征,将这两种不同层次的特征直接拼接是不匹配的。随着网络的深入,后续的跳接层所连接的编码器的特征相比于第一个跳接层经历了更多处理,层次更高,而对应的解码器的特征则层次更低,因此这两个对应特征的不匹配程度有所降低。

为了减小这种不匹配程度,本文对跳接层进行了改进,不直接将编码器的特征连接到解码器,而是添加一系列连续的卷积层对这些特征进行处理,提升特征的层次以增加和解码器对应特征的匹配度。对于所采用的卷积层个数,主要考虑两个因素:1)沿着U-Net网络向下(沿“U”型底部方向),跳接层两侧特征的不匹配度在下降,因此所需要的卷积层个数也要逐渐减小;2)跳接层两侧的特征所包含的信息本身会通过向下的通道进行流动,跳接层不需要完全弥补这一特征层次的差异,否则便失去了跳接的意义。因此,综合考虑,最终选择的卷积层个数为4、3、2、1,卷积核大小为 3×3 。此外,为了使学习过程更容易,在卷积层之外还引入了残差层,通过卷积核大小为 1×1 的卷积层实现,再与卷积层之后的结果相加。最终跳接层的网络结构如图1(b)所示,其中深色方块对应卷积核大小为 3×3 的卷积层,浅色方块对应卷积核大小为 1×1 的卷积层,即残差层。

2.3 损失函数及训练参数

用神经网络提升图像质量的目标是建立起低质量的输入图像 y_0 与高质量的参考图像 y^* 之间的映射,设计损失函数时要使得预测结果 y 与参考图像 y^* 之间的误差最小,一种较为经典的选择为均方差

差(MSE),但是MSE作为图像评价指标并非与人眼的视觉效果完全一致^[36],并且使用MSE作为损失函数所训练出的结果有过度平滑的趋势。结构相似性指数(SSIM)是更为符合人眼视觉的评价指标,综合考虑了图像的结构和纹理特征^[37]。评估两幅图像之间的SSIM需要先选择一定大小的窗口,对窗口内的图像计算局部相似性指数,然后滑动窗口遍历整幅图像取平均后得到完整的SSIM。预测结果中截取的图像窗口 m 与参考图像中截取的图像窗口 m^* 之间的SSIM为

$$R_{\text{SSIM}}(m, m^*) = \frac{2\mu_m \mu_{m^*} + c_1}{\mu_m^2 + \mu_{m^*}^2 + c_1} \cdot \frac{2\sigma_{mm^*} + c_2}{\sigma_m^2 + \sigma_{m^*}^2 + c_2} = l(m, m^*) \cdot d_{\text{cs}}(m, m^*), \quad (9)$$

式中: μ_m 和 μ_{m^*} 分别为 m 和 m^* 的平均值; σ_m^2 和 $\sigma_{m^*}^2$ 分别为 m 和 m^* 的方差; σ_{mm^*} 为 m 和 m^* 的协方差; c_1 和 c_2 为常数,目的是维持分式的稳定,避免分母为0,一般取 0.01^2 和 0.03^2 ;窗口大小一般取 $11 \text{ pixel} \times 11 \text{ pixel}$ 。SSIM的计算可以写成两项相乘的形式, $l(m, m^*)$ 比较两幅图像之间的亮度差异, $d_{\text{cs}}(m, m^*)$ 比较两幅图像之间的结构和对比度差异。

多尺度结构相似性指数(MS-SSIM)^[38]是SSIM的一种改进形式,可以融合不同分辨率的图像细节。对原图进行不同尺度 $j = \{1, \dots, M\}$ 的下采样,减小图像面积后计算尺度 M 下的 $l_M(m, m^*)$ 和多个不同尺度 j 下的 $d_{\text{cs}_j}(m, m^*)$,得到MS-SSIM为

$$R_{\text{MS-SSIM}}(m, m^*) = l_M(m, m^*) \cdot \prod_{j=1}^M d_{\text{cs}_j}(m, m^*). \quad (10)$$

下采样时图像缩放因子为2,当尺度 $M=1$ 时,图像尺寸不变;当尺度为 $M=2$ 时,图像宽度和高度均缩小一半。

MS-SSIM可以更好地反映预测结果与参考图像之间的结构和纹理特征,但是对绝对值偏差不敏感,因此采用绝对值误差 L_1 与MS-SSIM相结合的损失函数形式^[39]作为改进的U-Net网络的损失函数,

$$L = \alpha(1 - R_{\text{MS-SSIM}}) + (1 - \alpha)L_1, \quad (11)$$

式中: $L_1 = |y^* - y|$; α 为常数,一般取0.84。

网络训练时采用Adam算法^[40]对参数进行优化,以使损失函数达到最小值,初始学习率为 5×10^{-5} ,训练过程中Adam优化器会对学习速率进行调节,Adam中用到的一阶矩估计和二阶矩估计的指数衰

减率 β_1 和 β_2 分别为 0.9 和 0.999。

网络训练所用环境为 Tensorflow 2.0 框架, 所用图形处理器为 NVIDIA GeForce RTX 2080Ti, 显存大小为 11 GB。

3 实验结果

3.1 仿真数据集

为了验证改进的 U-Net 网络对稀疏视角下重建图像质量的提升效果, 本文首先通过数值仿体进行仿真实验。本文所用数据集为 Hoover 等^[41]采集的眼底成像数据集 DRIVE, 采用其中经由人工分割的血管图像作为数值仿体。DRIVE 数据集中共有 40 张图, 通过裁剪、旋转等操作将图片扩增至 1000 张, 每张图的像素大小为 256 pixel \times 256 pixel, 随后利用该 1000 张图像生成仿真的光声信号, 然后进行光声图像重建。光声仿真模拟直径为 80 mm 的环形换能器对信号的探测过程, 图像位于环形换能器中心位置, 图像所在区域大小为 25 mm \times 25 mm。

对于上述 1000 幅图像, 将其分成三部分: 训练

集 800 幅、验证集 100 幅、测试集 100 幅。网络训练的批大小为 8, 学习率为 5×10^{-5} , 总的训练轮次为 300, 从测试集中选取一组结果, 如图 2 所示。从图 2 可以看到, 在阵元数 $N=32$ 时稀疏迭代重建的图像效果较差, 有明显的条纹伪影, U-Net 和改进的 U-Net 训练得到的结果相比稀疏迭代重建的图像质量有明显的提升。由于作为网络输入的稀疏重建图像质量较差, 可以看到 U-Net 预测的结果尽管已经十分接近参考图像, 但仍然有一定差别, 粗血管内部均匀性较差, 改进的 U-Net 预测的结果则克服了这个问题, 图像质量更优, 得到了接近参考图像的结果。

对于阵元数 $N=128$ 情形下的结果, 难以用肉眼辨别不同方法重建结果的差异, 沿图 2(d) 和 (h) 中箭头所示区域取一轮廓线观察其幅值差异, 不同方法重建的图像中对应位置的幅值如图 3 所示。可以看出无论是 $N=32$ 还是 $N=128$, 相比于 U-Net, 改进 U-Net 的结果都具有与参考图像更为接近的轮廓。

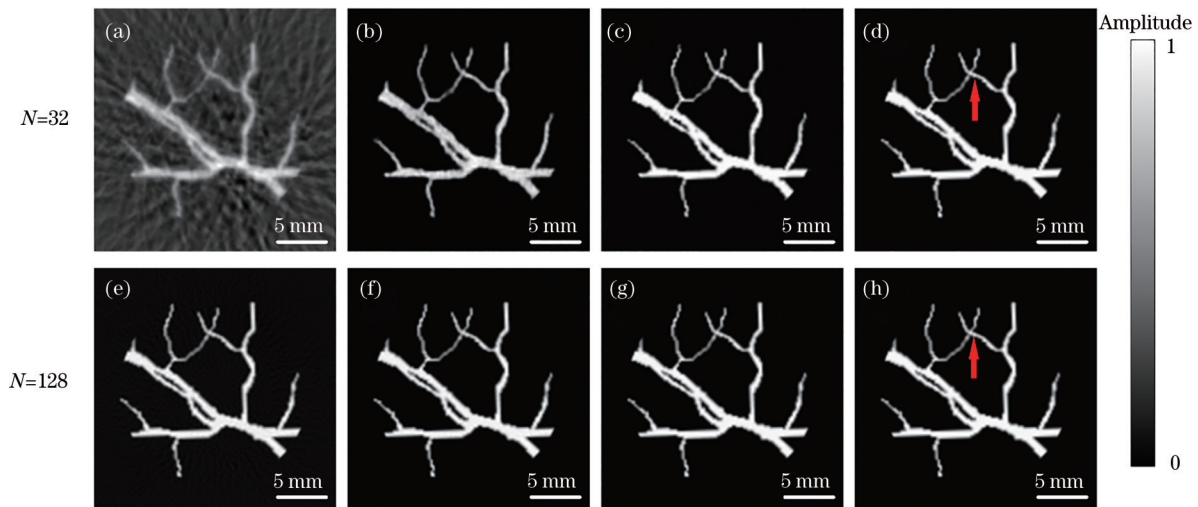


图 2 仿真数据集上的重建结果。(a)(e) 迭代重建图像; (b)(f) U-Net 网络输出结果; (c)(g) 改进 U-Net 网络输出结果; (d)(h) 参考图像

Fig. 2 Reconstruction results on simulation dataset. (a) (e) Images reconstructed by an iterative algorithm; (b) (f) U-Net outputs; (c) (g) modified U-Net outputs; (d) (h) reference images

对测试集中所有的 100 个样本, 统计预测结果的平均峰值信噪比 (PSNR) 和 SSIM, 如表 1 所示, 表中加粗数据为每种指标的最大值。两种网络的定量指标相对于传统迭代方法均有较大幅度提升, 阵元数 $N=32$ 时 PSNR 从 22.18 提升到了 31 以上, SSIM 从 0.231 提升到了 0.97 以上; 阵元数 $N=128$ 时 PSNR 从 40.53 提升到了 48 以上, SSIM 从 0.884 提升到了 0.99 以上。相比于 U-Net, 改进的 U-Net

的指标仍然有所提升, 阵元数 $N=32$ 时提升较为明显, 阵元数 $N=128$ 时 PSNR 有微弱提升, SSIM 指数持平。总体而言, 改进的 U-Net 相比于 U-Net 网络, 其结果在阵元数更少, 输入图像质量退化更为严重的情况下有较为明显的优势。

3.2 实验数据集

本文采用的活体成像数据集为实验采集到的小鼠横截面的光声图像, 实验所用换能器为阵元数

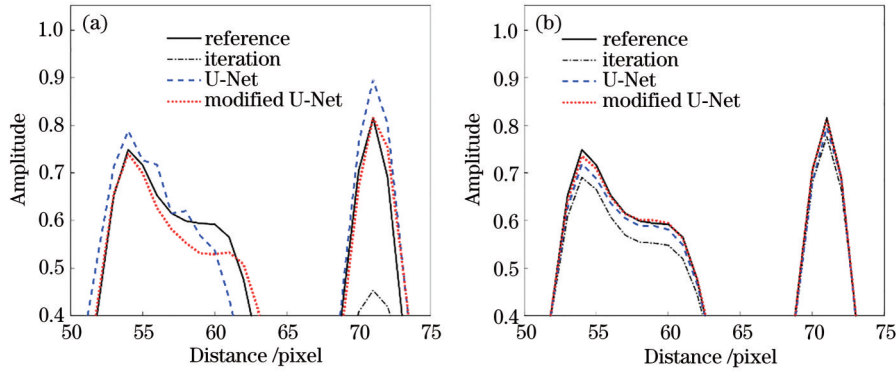


图 3 比较图 2(d)和(h)箭头标注处的幅值。(a) $N=32$;(b) $N=128$

Fig. 3 Comparison of profiles labeled by arrows in Fig. 2 (d) and (h). (a) $N=32$; (b) $N=128$

表 1 仿真数据集中不同方法的定量比较

Table 1 Quantitative comparisons of different methods on simulation dataset

| Item | Iteration | | U-Net | | Modified U-Net | |
|------|-----------|---------|--------|---------|----------------|--------------|
| | $N=32$ | $N=128$ | $N=32$ | $N=128$ | $N=32$ | $N=128$ |
| PSNR | 22.18 | 40.53 | 31.23 | 48.37 | 34.23 | 50.90 |
| SSIM | 0.231 | 0.884 | 0.975 | 0.996 | 0.987 | 0.996 |

512, 直径为 80 mm 的环形阵列, 成像区域位于环形换能器中心位置, 图像所在区域大小为 $25\text{ mm} \times 25\text{ mm}$ 。实验采集了 8 只小鼠的数据, 每只小鼠从头部到尾部以均匀间隔采集 128 张光声图像, 以阵元数 $N=512$ 情形下的重建结果作参考图像, 阵元数 $N=32$ 和 $N=128$ 情形下的重建图像作为训练和测试用的图像。实验中, 选取 6 只小鼠的光声重建图像作为训练集, 共 768 组, 1 只小鼠的光声图像作为验证集共 128 组, 1 只小鼠的光声重建图像作为测

试集共 128 组。

网络训练的批大小为 8, 学习率为 5×10^{-5} , 总的训练轮次为 200, 从测试集中选取两组不同截面的结果, 如图 4 和图 5 所示。图 4 为小鼠肝脏所在位置处截面的光声重建图像箭头 1 指示区域为腹部主动脉, 箭头 2 指示区域为肝脏内部的一根血管, 图 4(e) 中虚线所示为小鼠肝脏所在区域。在阵元数 $N=32$ 的情形下, 几乎无法从迭代重建图像中得到有效信息, 而经过 U-Net 处理的图像中可以看到腹部主动脉, 但是血管形状失真较严重, 改进的 U-Net 重建的图像则可以看到形状更接近参考图像的腹部主动脉截面。对于箭头 2 所指小血管, 两种网络重建的图像均无法得到, 主要是因为输入图像中的有效信息不足。在阵元数 $N=128$ 的情形下, 尽管输入图像存在伪影, 但是已经能够提供较多的细节, 两种网络重建的图像质量更高, 重建出的腹

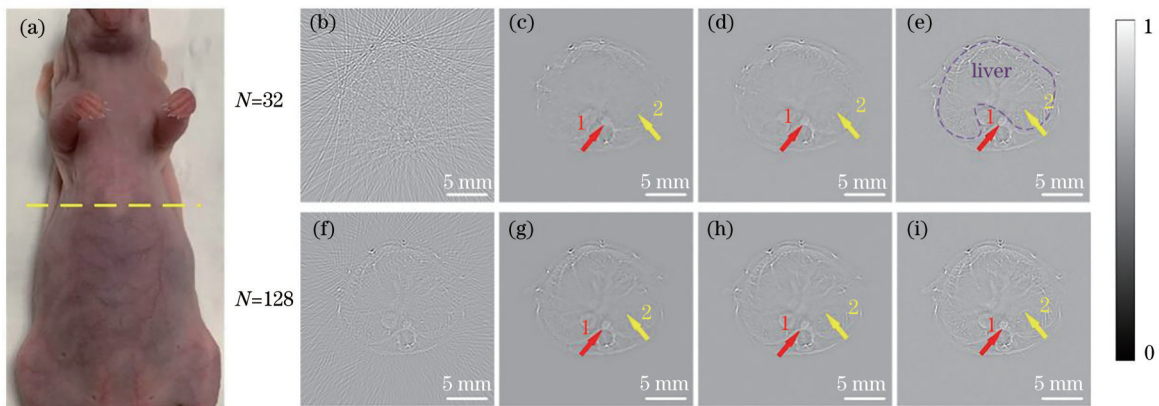


图 4 小鼠肝脏位置处实验数据集重建结果。(a) 成像位置示意; (b)-(f) 迭代重建图像; (c)-(g) U-Net 网络输出结果; (d)-(h) 改进 U-Net 网络输出结果; (e) 肝脏区域示意图; (i) 参考图像

Fig. 4 Reconstruction results of a cross section of the liver of a mouse on experimental dataset. (a) Position of imaging; (b)-(f) images reconstructed by an iterative algorithm; (c)-(g) U-Net outputs; (d)-(h) modified U-Net outputs; (e) demonstration of liver area; (i) reference image

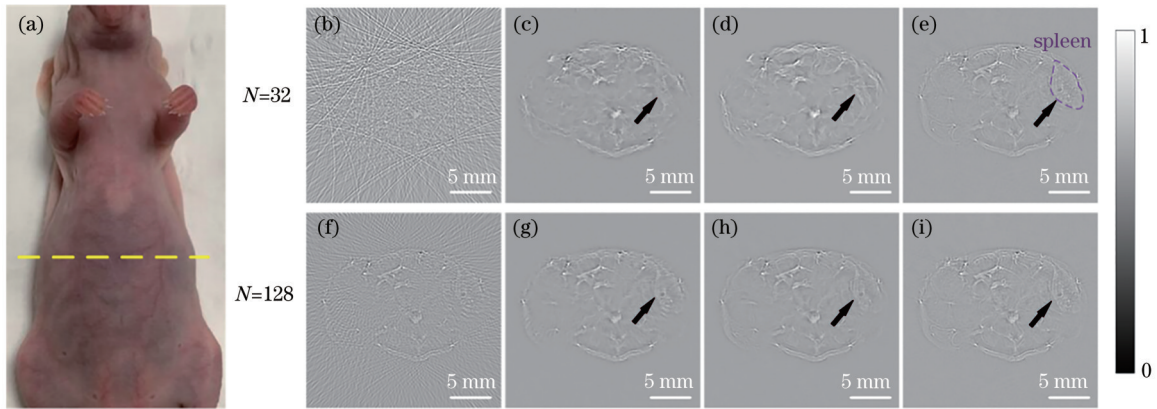


图 5 小鼠脾脏处实验数据集重建结果。(a)成像位置示意;(b)(f)迭代重建图像;(c)(g)U-Net网络输出结果;(d)(h)改进 U-Net网络输出结果;(e)脾脏区域示意图;(i)参考图像

Fig. 5 Reconstruction results of a cross section of the spleen of a mouse on experimental dataset. (a) Position of imaging; (b) (f) images reconstructed by an iterative algorithm; (c) (g) U-Net outputs; (d) (h) modified U-Net outputs; (e) demonstration of spleen area; (i) reference image

部主动脉都接近参考图像,也都重建出了箭头 2 所指的小血管,不同之处在于 U-Net 重建出的小血管隐约可见,改进的 U-Net 重建出的小血管则更为清晰。

图 5 为小鼠脾脏所在位置处的光声图像,箭头指示区域为脾脏,图 5(e)虚线所示为小鼠脾脏所在区域。在阵元数 $N=32$ 情形下,U-Net 和改进的 U-Net 重建得到的图像中都可以看到脾脏的轮廓,但 U-Net 重建得到的细节较为模糊,改进的 U-Net 重建得到的细节较为清晰。在阵元数 $N=128$ 情形下,U-Net 和改进的 U-Net 重建得到的图像都优于阵元数 $N=32$ 情形下的图像,并且改进的 U-Net 重建图像中脾脏轮廓更为清晰。

对测试集中所有的 100 个样本,统计预测结果的 SSIM 和 PSNR,如表 2 所示,表中加粗数据为每种指标的最大值。与仿真数据集上得到的结果类似,两种网络的定量指标相对于传统迭代方法均有较大提升,阵元数 $N=32$ 时 PSNR 从 20.03 提升到了 30 以上,SSIM 从 0.117 提升到了 0.7 以上;阵元数 $N=128$ 时 PSNR 从 26.47 提升到了 34 以上,SSIM 从 0.383 提升到了 0.8 以上。相比于 U-Net,

表 2 实验数据集中不同方法的定量比较

Table 2 Quantitative comparisons of different methods on experimental dataset

| Item | Iteration | | U-Net | | Modified U-Net | |
|------|-----------|---------|--------|---------|----------------|--------------|
| | $N=32$ | $N=128$ | $N=32$ | $N=128$ | $N=32$ | $N=128$ |
| PSNR | 20.03 | 26.47 | 30.08 | 34.74 | 31.17 | 35.16 |
| SSIM | 0.117 | 0.383 | 0.707 | 0.827 | 0.710 | 0.836 |

改进的 U-Net 增强图像的 PSNR 和 SSIM 指标都更好。

图 6 给出了以上仿真数据集和实验数据集中网络训练过程中的损失函数的变化曲线。网络训练过程中,由于仿真和实验数据集的输入图像质量相差比较大,仿真数据集训练的最终误差更小,需要更多的训练轮次寻找最优解,实验数据集则很快收敛。由于仿真数据集的训练集与验证集之间的数据分布并没有显著差异,因此训练过程中的训练损失与验证损失差异并不大;实验数据则不同,训练集与验证集采用的是来自不同小鼠的图像,数据之间有较大差异,训练过程中的训练损失与验证损失则有较大差距。并且, $N=32$ 时稀疏重建图像提供的有效信息更少,最终训练结果中验证损失与训练损失差距较大, $N=128$ 时情况有所好转,训练损失与验证损失差距缩小。

3.3 消融实验

仿真数据集和实验数据集的结果验证了改进的 U-Net 相比传统 U-Net 结构重建质量得到了提升,两种网络作训练时都采用了多尺度结构相似性指数 MS-SSIM 与绝对值误差 L_1 相结合的误差函数对网络进行训练,说明效果的提升来源于网络结构中跳接层的改进。为了分别比较跳接层和损失函数的作用,在换能器数量 $N=128$ 的数据集上进行消融实验,在原有的两组数据:U-Net+混合误差函数(L_{mix})和改进的 U-Net+混合误差函数(L_{mix})的基础上,增加一组改进的 U-Net+MSE 误差函数(L_{MSE}),结果如表 3 所示。

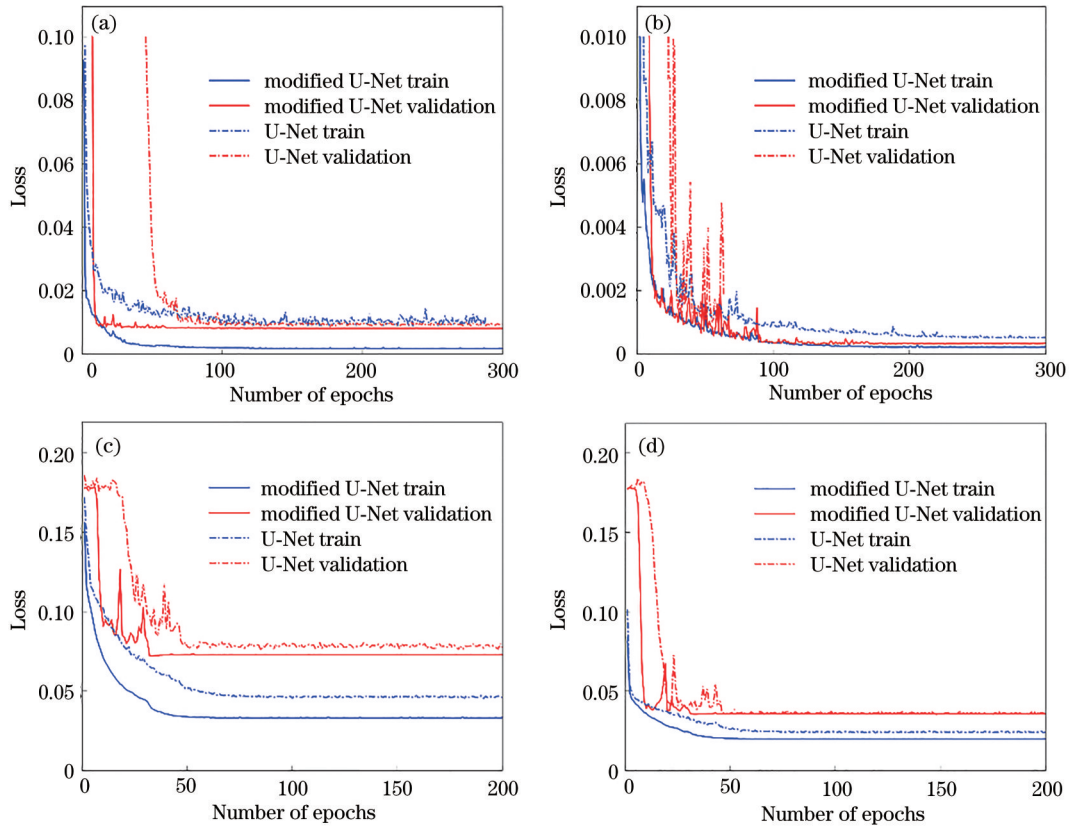


图 6 训练过程中的训练损失和验证损失比较。(a)仿真数据集,阵元数 $N=32$; (b)仿真数据集,阵元数 $N=128$; (c)实验数据集,阵元数 $N=32$; (d)实验数据集,阵元数 $N=128$

Fig. 6 Comparison of training losses and validating losses during training process. (a) Simulation dataset, element number $N=32$; (b) simulation dataset, element number $N=128$; (c) experimental dataset, element number $N=32$; (d) experimental dataset, element number $N=128$

表 3 消融实验结果

Table 3 Results of ablation experiment

| Item | U-Net+ L_{mix} | Modified U-Net+ L_{MSE} | Modified U-Net+ L_{mix} |
|------|------------------|---------------------------|---------------------------|
| PSNR | 34.74 | 35.13 | 35.16 |
| SSIM | 0.827 | 0.834 | 0.836 |

根据上表中消融实验结果,改进的 U-Net+ L_{mix} 的结果相比 U-Net+ L_{mix} 的结果有较明显提升,说明网络结构的改进有较明显的提升效果;改进的 U-Net+ L_{mix} 的结果相比改进的 U-Net+ L_{MSE} 的结果也有一定提升,但是提升幅度很小,说明损失函数也对网络训练带来了提升,但是相比于网络结构改进所带来的提升要小很多。可以得出结论,跳接层的引入和混合损失函数的采用都对结果起到了提升作用,并且作用主要来自跳接层的引入。

4 结 论

本文提出了一种基于神经网络的稀疏视角光

声重建图像质量提升方法,使用改进的 U-Net 网络对稀疏视角光声重建的图像做后处理,抑制稀疏视角引入的条纹伪影,提升细节表现能力。提升稀疏重建的质量,可以使光声成像系统利用较少的换能器数量重建出较高质量的图像,降低硬件成本的同时提高成像速度。

本文对 U-Net 网络结构的改进聚焦于跳接层所连接的两组对应特征的不匹配问题,引入新的跳接层,通过增加多层卷积处理操作,增加跳接层两侧特征的匹配度。仿真和实验数据结果表明,本文对 U-Net 所采用的改进策略可以提升从低质量的输入图像中提取有效信息的能力,进而增强传统 U-Net 网络的细节重建能力。改进的 U-Net 网络结构在换能器数量较少,原稀疏视角重建图像质量较低的情形下的图像质量提升尤为明显。论文提出的基于改进 U-Net 网络结构的稀疏视角下光声图像质量增强方法,对光声成像的性能提升和应用转化有重要意义。

参 考 文 献

- [1] Wang L V. Multiscale photoacoustic microscopy and computed tomography[J]. *Nature Photonics*, 2009, 3(9): 503-509.
- [2] Luke G P, Emelianov S Y. Label-free detection of lymph node metastases with US-guided functional photoacoustic imaging[J]. *Radiology*, 2015, 277(2): 435-442.
- [3] Tian C, Qian W, Shao X, et al. Plasmonic nanoparticles with quantitatively controlled bioconjugation for photoacoustic imaging of live cancer cells[J]. *Advanced Science*, 2016, 3(12): 1600237.
- [4] Tian C, Xie Z X, Fabiilli M L, et al. Imaging and sensing based on dual-pulse nonlinear photoacoustic contrast: a preliminary study on fatty liver[J]. *Optics Letters*, 2015, 40(10): 2253-2256.
- [5] Liu S D, Wang H, Zhang C X, et al. *In vivo* photoacoustic sentinel lymph node imaging using clinically-approved carbon nanoparticles[J]. *IEEE Transactions on Biomedical Engineering*, 2020, 67(7): 2033-2042.
- [6] Wang H, Liu S D, Wang T, et al. Three-dimensional interventional photoacoustic imaging for biopsy needle guidance with a linear array transducer[J]. *Journal of Biophotonics*, 2019, 12(12): e201900212.
- [7] Tian C, Zhang C X, Zhang H R, et al. Spatial resolution in photoacoustic computed tomography[J]. *Reports on Progress in Physics*, 2021, 84(3): 036701.
- [8] Tian C, Pei M L, Shen K, et al. Impact of system factors on the performance of photoacoustic tomography scanners[J]. *Physical Review Applied*, 2020, 13(1): 014001.
- [9] Xu M H, Wang L V. Universal back-projection algorithm for photoacoustic computed tomography [J]. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 2005, 71(1): 016706.
- [10] Wang K, Su R, Oraevsky A A, et al. Investigation of iterative image reconstruction in three-dimensional photoacoustic tomography[J]. *Physics in Medicine and Biology*, 2012, 57(17): 5399-5423.
- [11] Calvetti D, Morigi S, Reichel L, et al. Tikhonov regularization and the L-curve for large discrete ill-posed problems[J]. *Journal of Computational and Applied Mathematics*, 2000, 123(1/2): 423-446.
- [12] Shaw C B, Prakash J, Pramanik M, et al. Least squares QR-based decomposition provides an efficient way of computing optimal regularization parameter in photoacoustic tomography[J]. *Journal of Biomedical Optics*, 2013, 18(8): 080501.
- [13] Wang J, Wang Y Y. Photoacoustic imaging reconstruction using combined nonlocal patch and total-variation regularization for straight-line scanning [J]. *Biomedical Engineering Online*, 2018, 17(1): 105.
- [14] Zhang Y, Wang Y Y, Zhang C. Total variation based gradient descent algorithm for sparse-view photoacoustic image reconstruction[J]. *Ultrasonics*, 2012, 52(8): 1046-1055.
- [15] Li X P, Qi L, Zhang S Y, et al. Model-based photoacoustic tomography image reconstruction with non-local and sparsity regularizations[J]. *IEEE Access*, 2019, 7: 102136-102148.
- [16] Chen S S, Chen M H, Ma W F. Research on Automatic Classification of Optical Coherence Tomography Retina Image Based on Multi-Channel[J]. *Chinese Journal of Lasers*, 2021, 48(23): 2307001.
陈思思, 陈明惠, 马文飞. 基于多通道的光学相干层析成像视网膜图像自动分类研究[J]. *中国激光*, 2021, 48(23): 2307001.
- [17] Wang C, Qin H T, Lai G Y, et al. Automated classification of dual channel dental imaging of auto-fluorescence and white light by convolutional neural networks[J]. *Journal of Innovative Optical Health Sciences*, 2020, 13(4): 2050014.
- [18] He C E, Xu H J, Wang Z, et al. Automatic Segmentation Algorithm for Multimodal Magnetic Resonance-Based Brain Tumor Images[J]. *Acta Optica Sinica*, 2020, 40(6): 0610001.
何承恩, 徐慧君, 王忠, 等. 多模态磁共振脑肿瘤图像自动分割算法研究[J]. *光学学报*, 2020, 40(6): 0610001.
- [19] Zheng G, Jiang Y F, Shi C, et al. Deep learning algorithms to segment and quantify the choroidal thickness and vasculature in swept-source optical coherence tomography images[J]. *Journal of Innovative Optical Health Sciences*, 2021, 14(1): 2140002.
- [20] Zhang H L, Li Q, Guan X. An improved three-dimensional dual-path brain tumor image segmentation network[J]. *Acta Optica Sinica*, 2021, 41(3): 0310002.
张恒良, 李镛, 关欣. 一种改进的三维双路径脑肿瘤图像分割网络[J]. *光学学报*, 2021, 41(3): 0310002.
- [21] Zhu S Q, Wang J, Cai Y F. Low-Dose CT Denoising Algorithm Based on Improved Cycle GAN[J]. *Acta Optica Sinica*, 2020, 40(22): 2210002.
朱斯琪, 王珏, 蔡玉芳. 基于改进型循环一致性生成

- 对抗网络的低剂量 CT 去噪算法[J]. 光学学报, 2020, 40(22): 2210002..
- [22] Awasthi N, Jain G, Kalva S K, et al. Deep neural network-based sinogram super-resolution and bandwidth enhancement for limited-data photoacoustic tomography[J]. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2020, 67(12): 2660-2673.
- [23] Zhang J Y, Chen B, Zhou M, et al. Photoacoustic image classification and segmentation of breast cancer: a feasibility study[J]. IEEE Access, 2018, 7: 5457-5466.
- [24] Tong T, Huang W H, Wang K, et al. Domain transform network for photoacoustic tomography from limited-view and sparsely sampled data[J]. Photoacoustics, 2020, 19: 100190.
- [25] 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.
- [26] Davoudi N, Deán-Ben X L, Razansky D. Deep learning optoacoustic tomography with sparse data[J]. Nature Machine Intelligence, 2019, 1(10): 453-460.
- [27] Deng H D, Wang X H, Cai C J, et al. Machine-learning enhanced photoacoustic computed tomography in a limited view configuration[J]. Proceedings of SPIE, 2019, 11186: 111860J.
- [28] Lan H R, Yang C C, Jiang D H, et al. Deep learning approach to reconstruct the photoacoustic image using multi-frequency data[C]//2019 IEEE International Ultrasonics Symposium, October 6-9, 2019, Glasgow, UK. New York: IEEE Press, 2019: 487-489.
- [29] Shahid H, Khalid A, Liu X, et al. A deep learning approach for the photoacoustic tomography recovery from undersampled measurements[J]. Frontiers in Neuroscience, 2021, 15: 598693.
- [30] Feng J C, Deng J G, Li Z, et al. End-to-end Res-UNet based reconstruction algorithm for photoacoustic imaging[J]. Biomedical Optics Express, 2020, 11(9): 5321-5340.
- [31] Guan S, Khan A A, Sikdar S, et al. Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(2): 568-576.
- [32] Guan S, Khan A A, Sikdar S, et al. Limited-view and sparse photoacoustic tomography for neuroimaging with deep learning[J]. Scientific Reports, 2020, 10(1): 8510.
- [33] Liu P J, Zhang H Z, Lian W, et al. Multi-level wavelet convolutional neural networks[J]. IEEE Access, 2019, 7: 74973-74985.
- [34] Hariri A, Alipour K, Mantri Y, et al. Deep learning improves contrast in low-fluence photoacoustic imaging[J]. Biomedical Optics Express, 2020, 11(6): 3360-3373.
- [35] Aster R C, Borchers B, Thurber C H. Parameter estimation and inverse problems[M]. 3rd ed. Amsterdam: Elsevier, 2019: 165-170.
- [36] Zhang L, Zhang L, Mou X Q, et al. A comprehensive evaluation of full reference image quality assessment algorithms[C]//2012 19th IEEE International Conference on Image Processing, September 30-October 3, 2012, Orlando, FL, USA. New York: IEEE Press, 2012: 1477-1480.
- [37] 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: a Publication of the IEEE Signal Processing Society, 2004, 13(4): 600-612.
- [38] Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment[C]//The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003, November 9-12, 2003, Pacific Grove, CA, USA. New York: IEEE Press, 2003: 1398-1402.
- [39] 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.
- [40] Kingma D P, Ba J. Adam: a method for stochastic optimization[EB/OL]. (2014-12-22) [2021-02-04]. <https://arxiv.org/abs/1412.6980>.
- [41] Hoover A, Kouznetsova V, Goldbaum M. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response[J]. Proceedings. AMIA Symposium, 1998: 931-935.