arXiv:2305.01165v1 [eess.IV] 2 May 2023

基于自相似性的光声血管成像超分辨率：从手绘涂鸦中实现

YUANZHENG MA,[1] WANGTING ZHOU,[2] RUI MA,[3] SIHUAYANG,[3,4] YANSONG TANG,[1] 和 XUN GUAN[1,*]

[1]清华-伯克利深圳学院，清华大学深圳国际研究生院，清华大学，中国深圳，*518055*[2]教育部分子与神经影像工程研究中心，西安电子科技大学，中国陕西西安，*710126*[3]教育部激光生命科学重点实验室与激光生命科学研究所，生物光子学学院，华南师范大学，中国广州，*510631*[4]*yangsh@scnu.edu.cn*[*]*xun.guan@sz.tsinghua.edu.cn*

摘要： 基于深度学习的超分辨率光声血管成像（PAA）是一种强大的工具，可从欠采样图像中恢复血管图像，从而辅助疾病诊断。然而，由于训练样本稀缺，PAA超分辨率模型通常表现出不足的泛化能力，尤其是在连续监测任务中。为应对这一挑战，我们提出了一种新方法，采用利用伪造PAA图像训练的超分辨率PAA方法。首先，我们利用基于扩散的图像生成模型，根据手绘曲线生成逼真的人类嘴唇PAA图像。随后，我们使用这些伪造PAA图像训练基于自相似性的超分辨率模型。实验结果表明，在原域测试和跨域测试中，我们的方法均优于使用真实PAA图像训练的超分辨率模型。特别是，我们的方法利用深度学习模型伪造的图像提升了超分辨率重建的质量，表明尽管初始数据集有限，深度学习模型之间的协作仍可促进泛化。该方法为探索用于视觉任务的零样本学习神经网络展现出良好潜力。

© 2023 Optica出版集团

1. 引言

光声血管成像（PAA）是一种无标记血管成像方法，能够以高光学对比度揭示器官组织结构 [1–5]。进行氧饱和度测量、血流监测和癌性肿瘤诊断时 [6, 7], ，通常需要高空间分辨率的光声显微成像（PAA），以便对血管网络进行成像和测量。要实现高空间分辨率并观察体内的血管网络，需要较小的空间采样间隔。然而，高分辨率光声显微成像（PAA）通常需要较长的成像时间，从而降低成像速度。光声显微成像（PAA）极其耗时的成像过程严重阻碍了其临床应用，而对活体器官中的动态事件进行实时监测则更加困难重重 [8]。

一种新策略通过对低分辨率光声血管成像图像进行基于深度学习的超分辨率处理来提升其质量，从而加快成像过程 [9–11], 进而消除光声显微成像的速度瓶颈。然而，数据驱动的超分辨率模型通常需要大规模训练数据集，而收集这些数据往往耗时过长 [12]。例如，假设激光器-探测器-电机系统运行无故障，每张图像耗时 10 分钟 [13], ，那么建立一个包含 10240 张人类嘴唇光声显微成像图像、分辨率为 $10\mu m$ 的《小型》数据集需要 71.1 天；即使使用最快的基于微机电系统的采样系统，也仍需 7.11 天 [14]。尽管

数据集对于超分辨率模型学习低分辨率与高分辨率PAA图像之间的关系至关重要，但不同组织和器官所需的数据集也各不相同，进一步增加了数据集采集的工作量。

人们尝试基于真实PAA图像人为扩充训练数据集，从而减少实际采集的图像数量。[12] 引入迁移学习来补充训练图像，在不需要大规模实验数据集的情况下显著提升了图像质量。借助基于 Bicycle-GAN 的数据生成器，[14]实现了用于训练超分辨率 PAA 模型的 PAA 数据集扩充。一些研究人员采用基于 k-wave 的仿真方法，增加用于训练的高分辨率图像总量 [15–17]。尽管先前的方法试图减少需要采集的图像数量，却破坏了训练集与测试集之间的一致性，导致超分辨率模型在推理时表现不佳，且重建图像的可信度较低 [18, 19]。

本研究提出一种方法，只需使用数量显著较少的真实PAA图像即可实现 PAA 超分辨率。该方法包含两个阶段，每个阶段分别训练一个模型。第一阶段使用一小批真实PAA图像训练数据生成模型，该模型借助扩散概率模型，根据手绘涂鸦伪造PAA图像 [20]。基于扩散概率模型的方法具有更强的泛化能力：输入涂鸦或扩散过程的任何变化，都会在推理时生成一幅新图像；因此，该方法能够生成无限数量的伪造PAA图像。与图像旋转或翻转等常规数据增强技术不同，该阶段从有限的输入模式中学习，并提升训练数据的质量和多样性 [21]。第二阶段利用血管结构的自相似性，使用伪造图像训练 SwinIR，以实现 PAA 超分辨率 [22, 23]。定量分析表明，使用这些图像测试 PAA 超分辨率模型时，伪造PAA图像也能达到与真实PAA图像相当甚至更好的性能。这进一步验证了伪造图像的逼真度以及所提方法的有效性。为展示超分辨率模型对具有额外时间维度的数据的重建能力，我们通过重建 PAA 连续监测中人类嘴唇的视频帧序列，验证了模型的动态恢复效果。结果表明，该模型能够从低分辨率视频中实现时间一致的重建，为 PAA 在临床中的实时应用奠定了基础。

有兴趣的读者可前往 GitHub 上的图形用户界面，使用涂鸦生成血管图像：https://github.com/yuanzhengthu/handDrawnPAAImages。

2. 方法

2.1. 基于迭代强化的光声显微成像图像生成

用于将手绘涂鸦转换为相应PAA图像的图像生成模型基于超分辨率扩散概率模型。该模型以低分辨率PAA图像为条件，从高斯噪声图像 $y_T \sim \mathrm{N}(0, I)$ 迭代生成超分辨率血管图像。在扩散的每一步中，所学习的条件转移分布 $p_\theta(y_{t-1}|y_t, x) = \mathrm{N}_{x, f_\theta, t}$，$t \in [1, T]$，其中，$y_t$是第$t$ 步的噪声图像，f_θ 表示DeNN [20]，x 是低分辨率图像。随着 t 减小，生成的图像更有可能接近超分辨率图像 [20, 24]。使用与 [14], 相同的成像系统，我们收集了321张人类嘴唇的PAA图像，用于训练生成模型。如此规模的数据集可能不足以训练超分辨率模型，但足以让图像生成模型学习PAA图像的复杂结构和数据分布。在这种情况下，训练生成模型至少需要32张图像。我们的实验使用202张PAA图像训练生成模型，同时保留59张图像用于测试，60张用于验证。为了将涂鸦与逼真的血管图像 $\tilde{y}_0$联系起来，我们首先训练图像生成模型，使其学习低分辨率图像与

N倍放大的高分辨率PAA图像之间的关系（N 可以是2、4、8、16.....）。这些低分辨率图像是在OpenCV [25]中从高分辨率图像下采样得到的。训练完成后，生成模型能够根据低分辨率PAA图像推断出超分辨率PAA图像。然而，我们希望通过引入其他类型的低分辨率图像，生成更加逼真的高分辨率PAA图像。为此，我们绘制手绘涂鸦，并添加类似雨水的归一化高斯噪声，分别形成血管的主干和分支。随后，我们对这些图像进行下采样和上采样，以生成额外的低分辨率图像，并将其输入生成模型进行图像生成。

如图 1 所示，为了诱导模型将图 1(b) 中的手绘涂鸦识别为低分辨率光声图像，首先将涂鸦下采样至其 $\frac{1}{N}$ 尺寸，再上采样至原始尺寸。此处 N 可以是 2、4、8、16 或任意数值，但应与超分辨率扩散概率模型的放大倍率保持一致。在本实验中，我们设定 $N = 8$。随后，将类雨噪声 $\epsilon \mathrm{N}(0, I)$ 叠加其上，如图 1(a) 所示。再使用核 $[3, 3]$ 对涂鸦进行模糊处理，最终生成图像生成模型 x的输入，如图 1(c) 所示。为了避免逐段绘制涂鸦所需的大量工作，我们利用高分辨率图像的上采样子集作为血管的主体，从而加速光声显微成像图像的生成过程。如图 1 所示，我们随机裁剪一幅高分辨率图像（图 1 中标记为 HR），并将其上采样以生成图 1(b')。随后，将类雨噪声叠加到图 1(h') 上，如图 1(a') 所示，再使用高斯核对其进行模糊处理，得到图 1(c')。然后，将图 1(c') 输入生成模型，以生成逼真的高分辨率光声图像，如图 1(f') 所示。

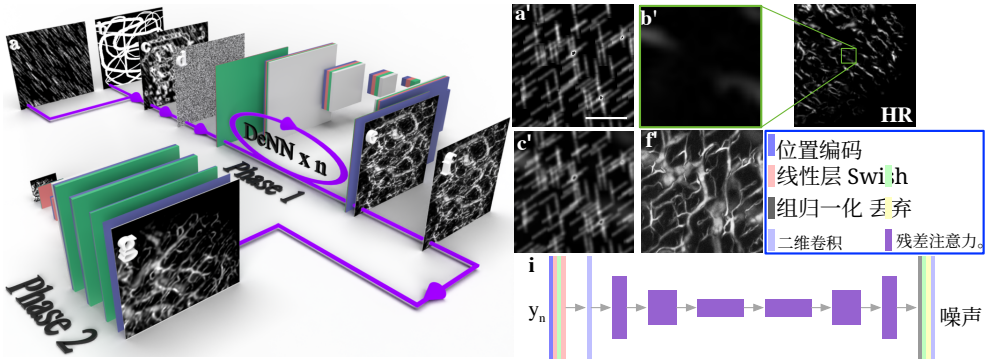


图 1. (a)-(g) 基于手绘涂鸦生成光声显微成像超分辨率血管图像的训练示意图。(a) 类雨噪声图像。(b) 手绘涂鸦。(c) 与 (b) 相关的图像生成模型输入，与 (a) 叠加。(d) 归一化高斯噪声图像。(e) 根据 (c) 生成的光声显微成像图像。(f) 归一化的光声显微成像图像。(g) 重建的超分辨率光声显微成像图像。(a')-(f') 基于裁剪图像生成光声显微成像超分辨率血管图像的训练示意图，对应左侧的 (a)-(f)。(a') 类雨噪声图像。(b') 高分辨率光声图像 **HR** 的随机子集。(c') (b') 与 (a') 叠加后的图像。(f') 通过图像生成模型根据 (c') 生成的光声显微成像图像。(i) 基于 UNet 的去噪神经网络 (DeNN) 的结构。**HR**：高分辨率光声显微成像图像。**ResAtten.**：基于 ResBlock 的注意力模块。比例尺：0.50 mm。

与 DeNN $\times$ n 相关的扩散过程通过基于 UNet 的去噪神经网络实现，如图 1(i) 所示。在推理过程中，将训练好的 DeNN 模型作为一系列级联链应用，逐步降低高斯噪声图像 y_T 中的噪声。该过程由马尔可夫链描述：在输入 x 的条件下，图像中的高斯噪声逐渐减少，从而生成逼真的光声显微成像图像 $\tilde{y_0}$ [9,26,27]。由于输入涂鸦中含有重叠的类雨噪声，通过修改扩散过程中的涂鸦或类雨噪声分布，可以生成多种血管图像

可通过在扩散过程期间修改涂鸦或类雨噪声的分布来生成。图像重建过程可描述如下：

$$y_{t-1} \leftarrow \frac{1}{\sqrt{\alpha_t}}(y_t - \frac{1-\alpha_t}{\sqrt{1-\gamma_t}} f_\theta(x, y_t, \gamma_t)) + \sqrt{1-\alpha_t}\epsilon_t, \tag{1}$$

其中， y_t 表示第 t 步的噪声图像， y_0 表示对应的高分辨率光声图像。超参数 α_t 由一个递增函数描述（详情见代码），且$\gamma_t = \prod_{i=1}^{t} \alpha_i$。变量 $\epsilon_t \sim \mathrm{N}(0, \mathbf{I})$，而 $f_\theta(x, y_t, \gamma_t)$表示DeNN估计出的图像噪声。在此过程中，我们将加噪过程视为把DeNN的输出加到图像上，并将去噪过程视为从图像中减去DeNN的结果。通过逐步将图像退化为高斯噪声，可以在加噪过程中训练DeNN。这样，在推理阶段，逆转加噪过程即可通过去噪过程从涂鸦生成逼真的PAA图像。

2.2. 使用*K-means*自动筛选高峰值信噪比（*PSNR*）光声图像

由于生成的 PAA 并非都具有较高的峰值信噪比，因此采用基于 K 均值的筛选方法，选出峰值信噪比较高的 PAA图像 [28]。峰值信噪比较低的图像具有噪声背景，这意味着低峰值信噪比图像中的低强度像素数量远多于高峰值信噪比图像。如图 2 所示，人工从生成的 PAA图像中分别选出 10 张峰值信噪比较高和较低的图像，作为优质和劣质标准。随后，分别计算这些图像的 256 级直方图：高峰值信噪比 PAA图像的直方图为 $H_g = (\mathrm{h}_1, \mathrm{h}_2, \mathrm{h}_3, \mathrm{h}_4, ..., \mathrm{h}_{256})$，低峰值信噪比 PAA图像的直方图为 $H_b = (l_1, l_2, l_3, l_4, ..., l_{256})$。输入光声图像 I_n 的 256 级直方图计数为 $H_n = (x_n{,}1, x_n{,}2, x_n{,}3, x_n{,}4, ..., x_n{,}256)$。随后，分别计算 H_n 与 H_g 之间以及H_n 与 H_b 之间的距离。

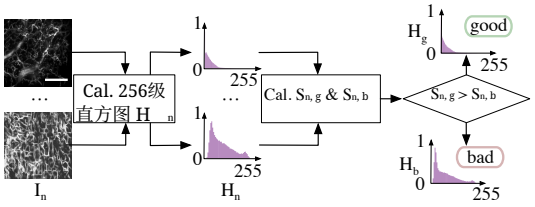


图 2. 用于筛选 PAA图像的 K 均值聚类算法流程图。 I_n：n_{th} 输入光声图像。 H_n：I_n的直方图。 $H_g(H_b)$：高峰值信噪比（低峰值信噪比）标准的直方图， $S_g(S_b)$表示输入 I_n 与优质标准（劣质标准）之间的相似度。比例尺：0.50mm。

如图 2 所示，根据输入图像 I_n 与高峰值信噪比标准的相似度对其进行分类。输入图像与高峰值信噪比标准之间的相似度 $S_{n,g}$ 的格式如下

$$S_{n,g} = \frac{1}{\sqrt{\sum_{i=1}^{256} |x_{n,i} - h_i|^2}}. \tag{2}$$

同样，$S_{n,b}$ 以相同的方式计算。如流程图所示，根据输入图像与其中一组的接近程度，将其分类为 $good$ 或 bad。

2.3. 用于 *PAA* 重建的基于自相似性的超分辨率模型

利用生成的高峰值信噪比PAA图像，我们有效地训练超分辨率模型，以提升血管图像的质量和分辨率。已有研究充分记录了血管的分形结构 [29, 30], ，通过将浅层和深层特征提取与SwinIR相结合，我们能够利用这种自相似性来提高血管图像的重建质量 [22,23]。在超分辨率模型中加入残差跳跃连接，可以进一步提升重建质量 [22]。这种方法对PAA尤为重要，因为PAA能够对生物体的精细血管结构进行成像。首先，我们使用上一阶段生成的PAA图像训练SwinIR，并默认采用随机翻转和旋转等数据增强技术。我们使用重建图像与相应真实值之间的L1距离来评估模型的性能。超分辨率PAA重建所采用的损失函数通过实验选定如下

$$\mathcal{L} = ||I_{SR} - I_{HR}||_1 + \alpha||w||_1, \tag{3}$$

其中， I_{SR} 表示重建得到的超分辨率光声显微成像图像， I_{HR} 表示真实值， α 表示用于调整 L1 损失与正则化项之间比例的超参数， w 表示超分辨率模型的权重向量。在 4 倍超分辨率任务中，我们选择 $\alpha = 10^{-4}$ ；在 8 倍超分辨率任务中，我们选择 $\alpha = 10^{-3}$ 。

3. 结果与讨论

3.1. 伪造与真实*PAA*图像的区分

如图3所示，图3(a)中的图像及其下方的图像为真实的人类嘴唇PAA图像，图3(b)中的图像及其下方的图像则为伪造图像。我们邀请了14位光声成像专家，使用虚线方框中所示的流程图区分真实图像和伪造图像。我们采用随机裁剪和随机打乱的方法，创建了一个由40张PAA图像组成的数据集，其中既有真实图像，也有伪造图像。随后，我们将该数据集发送给14位专家，请他们区分真实图像和伪造图像。如图3(c)所示，标签**Y** 和**N** 分别表示图像被归类为真实PAA图像或伪造PAA图像。每个红点或黑叉代表一位专家的判断结果，y轴表示被归类为**Y** 或**N**的样本数量。结果显示，真实的人类嘴唇PAA图像被判定为真实图像的概率为48.04%，而伪造图像被归入真实图像组的概率为46.79%。基于迭代精炼的PAA图像生成方法的欺骗率为49.375%，远高于增强型超分辨率生成对抗网络的欺骗率，后者仅为20% [15]。

为了以高分辨率PAA图像为标准，对伪造PAA图像与其他图像进行定量比较，我们使用了弗雷歇 *Inception* 距离 [31], ，该指标根据*Inception V3*模型的均值和协方差矩阵来衡量两组图像之间的距离。尽管*Inception V3* 模型并未在真实PAA图像上进行训练，但仍可使用弗雷歇 *Inception* 距离区分不同的血管图像集。伪造PAA图像与真实PAA图像之间的弗雷歇 *Inception* 距离为

$$\mathcal{FID}(I, G) = ||\mu_I - \mu_G||_2^2 + Tr(\Sigma_I + \Sigma_G - 2(\Sigma_I\Sigma_G)^{\frac{1}{2}}), \tag{4}$$

其中， I、 G 分别表示输入图像和真实值， μ_I、 μ_G 是输入图像和真实值的均值矩阵， $Tr()$是矩阵的迹， Σ_I、 Σ_G 分别是输入图像和真实值的协方差矩阵。不同数据集之间的弗雷歇 *Inception* 距离结果列于图3(d)中，显示真实PAA图像与伪造PAA图像之间的距离远小于其与其他图像之间的距离。真实PAA

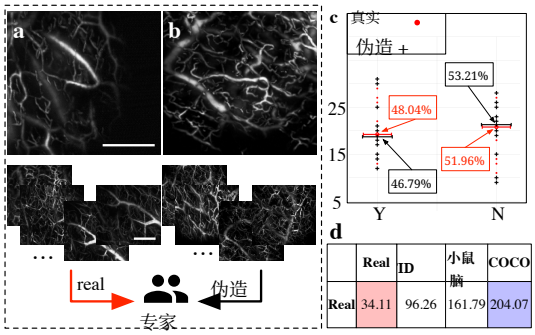


图3. 左侧虚线框圈出了14位光声成像专家用于区分真实与伪造图像的方法。(a) 人类嘴唇的真实光声显微成像图像。(b) 人类嘴唇的伪造光声显微成像图像。(c) 14位光声成像专家将真实和伪造PAA图像分类为**Y** （判定为真实PAA图像）和**N** （判定为伪造PAA图像）的散点图。(d) 以人类嘴唇的真实PAA图像为标准，比较人类嘴唇的真实PAA图像（**真实**）、人类嘴唇的伪造PAA图像（**伪造**）、小鼠脑的PAA图像（**小鼠脑**）和COCO 数据集（**COCO**）的弗雷歇 *Inception* 距离分数。最长距离以蓝色标出，最短距离以红色标出。比例尺：0.50 mm。

图像集与10幅手动选择的PAA图像之间的距离为34.11。这是因为10幅手动选择的PAA图像无法充分代表整个数据集。显然，与小鼠脑和COCO相比，伪造图像与真实图像更加相似，这一点可由较小的FIR 值看出。这些伪造PAA图像通过基于 k 均值的方法进行筛选，并收集用于超分辨率训练。

3.2. 对欠采样*PAA*图像进行超分辨率重建的评估

如图4所示，我们通过测量所提方法在两个变量下的峰值信噪比（PSNR）和结构相似性指数（SSIM），量化其泛化能力。在实验中，我们展示了超分辨率模型能够从图像的4倍下采样版本中重建高质量血管图像。

首先，我们在 [8, 32, 64, 128, 202] 范围内改变用于训练图像生成模型的真实图像数量。使用所有这些图像生成模型伪造1281张图像，以训练超分辨率模型；同时，根据真实图像数量不同，计算相应模型的SSIM和PSNR。作为对照，我们也使用相同数量的真实图像，在不进行图像伪造的情况下直接训练超分辨率模型。如图4(g)和(i)所示，无论是否使用图像伪造模型，PSNR和SSIM都会随着真实图像数量的增加而提高。根据 [32], ，当模型具有足够的特征提取参数时，规模大且质量好的数据集能够带来更好的结果。在我们的实验中，基于伪造图像进行超分辨率重建所得的PSNR始终高于基于真实图像的结果。模型性能的提升归因于生成模型能够扩充有限的输入模式，增加用于训练的高质量PAA图像数量 [21]。然而，当仅有8张真实PAA图像时，图像伪造的SSIM较低。虽然数据集规模为8不足以训练生成模型，但当数据集规模增加到32时，训练便成为可行。

随后，我们将图像生成模型训练中真实图像的数量固定为64，并调整生成的伪造图像数量。结果如图4(h)和(j)所示。随着伪造图像数量从32增加到1281，峰值信噪比和结构相似性指数持续提高。此外，在将生成的PAA图像纳入训练

数据集之前，我们会对其进行筛选，以确保质量。该方法有助于增加用于训练的高质量PAA图像数量，从而改善重建结果。

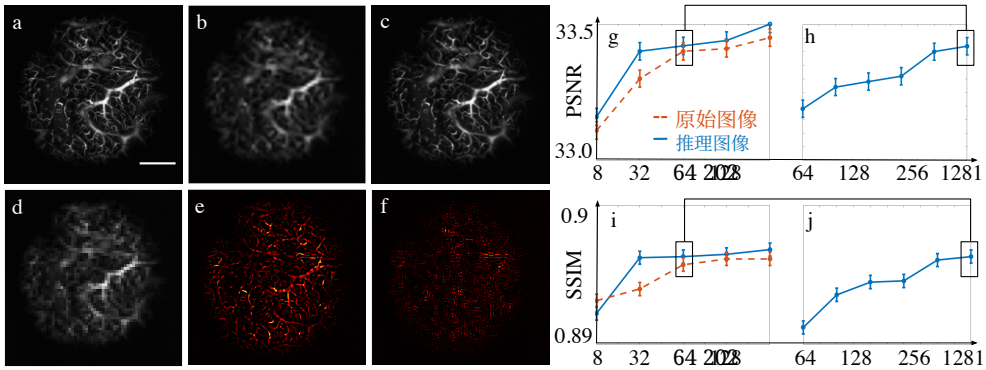
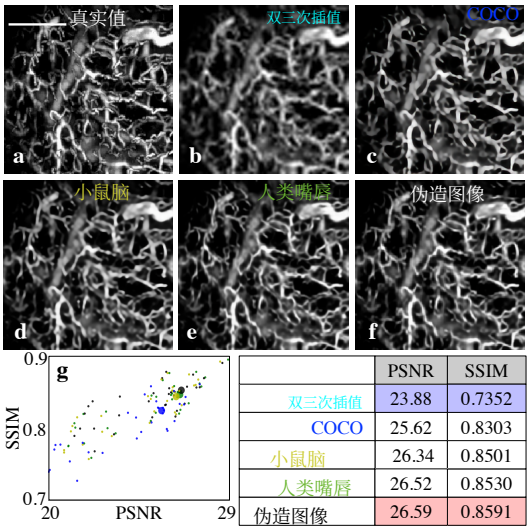


图4. 使用不同数量真实图像训练的超分辨率模型重建能力比较。(a) 真实值图像。(b) 使用双三次算法重建的高分辨率图像。(c) 使用我们的方法从202张真实图像中训练后重建的超分辨率图像。(d) 低分辨率图像。(e) 双三次插值重建图像与真实值之间的差异。(f) 使用我们的方法通过202张真实图像重建的图像与真实值之间的差异。(g)(i) 使用1281张伪造PAA图像训练的超分辨率模型的峰值信噪比和结构相似性指数；其伪造生成模型使用数量不等的真实图像训练，数量从8张到202张不等。(h)(j) 使用数量不等的伪造PAA图像训练的超分辨率模型的峰值信噪比和结构相似性指数；其伪造生成模型使用64张真实图像训练。(g)(h)中的峰值信噪比数据点以及(i)(j)中的结构相似性指数数据点，分别通过连线框标示相同的数据点。比例尺：0.50毫米。

3.3. 在其他人类嘴唇光声显微成像数据集上测试欠采样图像的超分辨率重建

为进一步考察模型从不同数据集中重建人类嘴唇光声显微成像图像的泛化能力，并评估其泛化性能，我们从 [14]，中随机选取 30 张光声显微成像图像，并使用不同训练数据集对其进行超分辨率重建，包括 COCO 数据集 [33], 中的小鼠脑光声显微成像图像 [8], , 以及真实的人类嘴唇光声显微成像图像。在此，我们使用 13277 张生成人类嘴唇的图像和 202 张真实图像来训练超分辨率模型。尽管超分辨率模型和训练超参数相同，图 5(a)-(f) 之间仍存在明显差异，这凸显了数据集一致性对于光声显微成像超分辨率的重要性。值得注意的是，伪造的人类嘴唇图像的视觉重建效果不逊于真实光声显微成像图像，这进一步验证了使用伪造图像训练的超分辨率模型的泛化能力。

图5(g)中的散点图和表格量化了这种泛化能力。散点图展示了从不同数据集中重建的30幅PAA图像的质量指标，图中以不同颜色的圆点标示，这些颜色与表格中的文字颜色相对应。不同数据集的均值以比单幅图像对应圆点更大的彩色圆点突出显示。由于双三次插值所得结果无法与其他方法所得结果进行比较，因此未将其纳入该散点图。结果表明，使用人类嘴唇的伪造PAA图像训练的模型优于



	PSNR	SSIM
双三次插值	23.88	0.7352
COCO	25.62	0.8303
小鼠脑	26.34	0.8501
人类嘴唇	26.52	0.8530
伪造图像	26.59	0.8591

图 5. 不同重建方法的比较。(a) 人类嘴唇的高分辨率图像。(b) 采用双三次算法重建的图像。(c)-(f) 使用以下数据集训练的超分辨率模型重建的结果：(c) COCO 数据集；(d) 小鼠背部 PAA 图像；(e) 真实人类嘴唇 PAA 图像；(f) 伪造的人类嘴唇 PAA 图像。(g) 不同重建结果的峰值信噪比和结构相似性指数，均值以较大的圆圈标示；右下方的图表展示了不同重建方法所得结果的峰值信噪比和结构相似性指数，不同数据集以不同颜色标示。比例尺：0.50 mm。

使用其他数据集训练的模型在测试集上的表现，甚至优于使用真实PAA图像训练的模型。这主要是因为生成模型扩充了数据集，增加了血管图像的多样性。此外，经过筛选后，伪造PAA数据集包含超过10,000幅图像，远大于仅包含202幅图像的真实PAA数据集。值得注意的是，图5表格中的峰值信噪比和结构相似性指数指标低于图4(g)-(h)所示的指标，这是因为两个数据集的采样条件存在不可控的差异，例如激光强度、光斑直径和采样间隔不同。

我们的实验证实，所提出的方法具备重建器官图像的能力，尤其适用于训练数据集不足的情况。在临床环境中，光声显微成像可能使器官暴露于高能量激光下，从而造成损伤。然而，我们提出的方法能够提高深部器官图像的分辨率，同时最大限度地减少激光照射。

3.4. 对具有视频帧率的光声显微成像图像序列进行评估

体内动态监测需要以较快的成像速度获取高分辨率光声图像，而传统的基于MEMS的光声显微成像方法无法实现这一点。为展示相邻帧之间的连贯性（这是动态监测的关键），我们进一步重建了随时间变化的光声显微成像图像。为加快采样过程，我们直接记录器官的最大投影图像。高分辨率PAA图像的成像速度从1 180 帧/秒提升至1 4 帧/秒。

为了以视频帧率获取低分辨率光声血管成像图像，我们对包含33帧的原始视频进行空间下采样，下采样比例为1/4，与超分辨率模型的放大倍率相对应。随后，我们采用上述

方法重建动态监测视频。如图6所示，绿色方框标出了真实值、使用小鼠脑数据集训练的超分辨率模型的结果以及使用伪造人类嘴唇数据集训练的模型结果之间的差异；使用伪造人类嘴唇数据集训练的模型结果优于其他结果。此外，使用伪造的人类嘴唇图像训练的超分辨率模型，其结果比高分辨率序列的背景更干净，这意味着在欠采样重建过程中，图像噪声也被去除了。理论上，重建后的视频可将器官的监测速度提高16倍，因为在两个维度上，每个维度仅采集了四分之一的采样点。这证实了超分辨率消除了光声显微成像的成像速度瓶颈。

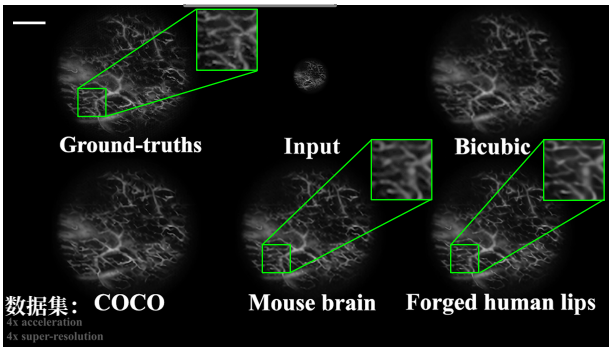


图6. 人类嘴唇欠采样动态光声显微成像的重建帧。比例尺：0.50mm（见可视化1）。

4. 结论

基于深度学习的超分辨率方法能够在保持光声显微成像时间分辨率的同时提高其空间分辨率。然而，现有的大多数光声显微成像超分辨率模型都需要使用规模庞大的数据集进行训练，其中至少包含数千张真实血管图像，而采集这些图像需要耗费极长时间。除了繁重的临床工作外，长时间的激光照射也会损伤目标器官。为克服这一限制，我们使用手绘血管图像训练超分辨率模型，并取得了优于其他模型的结果，包括优于使用真实PAA图像训练的模型。该方法减轻了基于扫描的图像采集所带来的过重工作负担，并拓展了光声显微成像在临床中的应用，使其能够实现快速高分辨率PAA成像，甚至组织实时视频成像。此外，我们的研究结果表明，使用生成模型生成的伪造图像训练的超分辨率模型优于使用真实PAA图像训练的模型。这说明迭代训练能够显著增强模型生成高度逼真、与真实图像极为相似的图像的能力，使其成为在缺乏或训练数据不足的情况下探索超分辨率技术的有前景工具。

资助。 深圳市科技创新委员会：基础研究计划一般项目（JCYJ20220530142809022）；稳定支持（WDZC20220811170401001）。国家重点研发计划（2021YFA1001000）。

披露声明。 作者声明不存在利益冲突。

数据可用性声明。 本文所报告结果的相关数据目前尚未公开，但可在提出合理请求后向作者获取。

代码可用性。 与本文相关的代码、模型及其他测试材料均可在线获取，并可从 https://github.com/yuanzhengthu/handDrawnPAAImages 下载。

参考文献

1. S. Hu、B. Rao、K. Maslov 和 L. V. Wang，“无标记光声眼科血管造影”，《光学快报》**35**，1–3（2010年）。2. C. Errico、J. Pierre、S. Pezet、Y. Desailly、Z. Lenkei、O. Couture 和 M. Tanter，“用于深层超分辨率血管成像的超快超声定位显微术”，《自然》**527**，499–502（2015年）。3. P. H. Reza、K. Bell、W. Shi、J. Shapiro 和 R. J. Zemp，“深度非接触式光声初始压力成像”，《光学》**5**，814（2018年）。4. L. Nie、S. Wang、X. Wang、P. Rong、Y. Ma、G. Liu、P. Huang、G. Lu 和 X. Chen，“利用靶向等离激元纳米星进行体内三维光声分子血管造影和治疗监测”，《小型》**10**，1585–1593（2014年）。5. L. Lin 和 L. V. Wang，“光声成像在临床肿瘤学中的新兴作用”，《自然综述：临床肿瘤学》**19**，365–384（2022年）。6. L. V. Wang、X. Wang、G. Ku、X. Xie 和 G. Stoica，“高分辨率光声断层成像”，《激光与光电子学会年会会议录》**2**，767–768（2004年）。7. F. Gao、X. Feng、R. Zhang、S. Liu、R. Ding、R. Kishor 和 Y. Zheng，“单个激光脉冲产生双光声信号，用于差分对比光声成像”，《科学报告》**7**，1–12（2017年）。8. A. DiSpirito、D. Li、T. Vu、M. Chen、D. Zhang、J. Luo、R. Horstmeyer 和 J. Yao，“利用深度学习重建欠采样光声显微图像”，《IEEE医学影像汇刊》**40**，562–570（2020年）。9. J. Li、C. Wang、T. Chen、T. Lu、S. Li、B. Sun、F. Gao 和 V. Ntziachristos，“在缺少标注实验数据的情况下，对深层组织进行基于深度学习的定量光声断层成像”，《光学》**9**，32（2022年）。10. C. Yang、H. Lan、F. Gao 和 F. Gao，“用于光声成像的深度学习综述”（2021年）。11. N. Davoudi、X. L. Deán-Ben 和 D. Razansky，“利用稀疏数据进行深度学习光声断层成像”，《自然·机器智能》**1**，453–460（2019年）。12. G. Godefroy、B. Arnal 和 E. Bossy，“利用可提供预测不确定性的深度学习方法补偿光声成像中的可见性伪影”，《光声学》**21**，100218（2021年）。13. W. Zhou、Z. Chen、Q. Zhou 和 D. Xing，“通过非接触式光声成像和光学相干断层成像对黑色素瘤及基底细胞癌进展进行光学活检：体内多参数表征肿瘤微环境”，《IEEE医学影像汇刊》**39**，1967–1974（2019年）。14. Y. Ma、K. Xiong、X. Hou、W. Zhang、X. Chen、L. Li 和 S. Yang，“用于少样本超分辨率光声血管造影的级联神经近似”，《应用物理快报》**121**（2022年）。15. A. Hauptmann、F. Lucka、M. Betcke、N. Huynh、J. Adler、B. Cox、P. Beard、S. Ourselin 和 S. Arridge，“基于模型的学习：加速有限视角三维光声断层成像”，《IEEE医学影像汇刊》**37**，1382–1393（2018年）。16. A. Yazdani、S. Agrawal、K. Johnstonbaugh、S. R. Kothapalli 和 V. Monga，“用于光声目标定位的同步去噪与定位网络”，《IEEE医学影像汇刊》**40**，2367–2379（2021年）。17. T. Tong、W. Huang、K. Wang、Z. He、L. Yin、X. Yang、S. Zhang 和 J. Tian，“用于有限视角和稀疏采样数据光声断层成像的域变换网络”，《光声学》**19**，100190（2020年）。18. N. Awasthi、G. Jain、S. K. Kalva、M. Pramanik 和 P. K. Yalavarthy，“基于深度神经网络的正弦图超分辨率与带宽增强：用于有限数据光声断层成像”，《IEEE超声、铁电与频率控制汇刊》**67**，2660–2673（2020年）。19. K. T. Hsu、S. Guan 和 P. V. Chitnis，“比较用于光声断层成像重建的深度学习框架”，《光声学》**23**，100271（2021年）。20. C. Saharia、J. Ho、W. Chan、T. Salimans、D. J. Fleet 和 M. Norouzi，“通过迭代精炼实现图像超分辨率”，《IEEE模式分析与机器智能汇刊》**PP**，1–14（2021年）。21. H. Kobeissi、S. Mohammadzadeh 和 E. Lejeune，“利用基于生成模型的增强训练数据集改进力学元模型”，《生物力学工程杂志》**144**，121002（2022年）。22. J. Liang、J. Cao、G. Sun、K. Zhang、L. Van Gool 和 R. Timofte，“SwinIR：使用 Swin Transformer 进行图像复原”，载于*IEEE*国际计算机视觉会议论文集，第2021年10月卷（2021年），第1833–1844页。23. D. Zhang、F. Huang、S. Liu、X. Wang 和 Z. Jin，“Swinfir：结合快速傅里叶卷积和改进训练，重新审视 SwinIR 图像超分辨率”（2022年）。24. J. Ho、A. Jain 和 P. Abbeel，“去噪扩散概率模型”，《神经信息处理系统进展》**2020年12月**，1–25（2020年）。25. G. Bradski，“OpenCV库”，《Dr. Dobb’s杂志：专业程序员软件工具》**25**，120–123（2000年）。26. J. Sohl-Dickstein、E. A. Weiss、N. Maheswaranathan 和 S. Ganguli，“利用非平衡热力学进行深度无监督学习”，载于第*32*届国际机器学习会议，*ICML 2015*，第3卷（2015年），第2246–2255页。27. H. Li、Y. Yang、M. Chang、S. Chen、H. Feng、Z. Xu、Q. Li 和 Y. Chen，“SRDiff：基于扩散概率模型的单幅图像超分辨率”，《神经计算》**479**，47–59（2022年）。28. J. Zhang、B. Chen、M. Zhou、H. Lan 和 F. Gao，“乳腺癌光声图像分类与分割：一项可行性研究”，《IEEE开放获取》**7**，5457–5466（2019年）。29. K. L. Uahabi 和 M. Atounti，“分形在医学中的应用”，《克拉约瓦大学年刊：数学与计算机科学系列》**42**，167–174（2015年）。30. E. Gabryś、M. Rybaczuk 和 A. Kędzia，“利用循环系统的分形模型进行血流模拟”，

《混沌、孤子与分形》**27**, 1–7（2006）。31. A. Borji，“生成对抗网络评估指标的优缺点：新进展”，《计算机视觉与图像理解》**215**, 103329 2022年( )。32. A. Ajiboye、R. Abdullah-Arshah、H. Qin 和 H. Isah-Kebbe，“使用……评估数据集规模对预测模型的影响”，《国际软件工程与计算机科学期刊》**1**, 75–84（2015年）。33. T. Lin、M. Maire、S. J. Belongie、L. D. Bourdev、R. B. Girshick、J. Hays、P. Perona、D. Ramanan、P. Dollár 和 C. L. Zitnick，“微软 COCO：上下文中的常见物体”，CoRR **abs/1405.0312**（2014年）。