

光声计算层析成像中的皮肤智能去除方法

袁怡鑫¹, 陈涛², 刘成波^{2**}, 孟静^{1*}¹曲阜师范大学计算机学院, 山东 日照 276826;²中国科学院深圳先进技术研究院生物医学与健康工程研究所, 广东 深圳 518055

摘要 光声计算层析成像(PACT)不需要外源性对比剂便可获取厘米级深度的光声图像。然而,来自皮肤的高强度光声信号遮盖了皮下深层组织信息,阻碍了感兴趣区域光声图像的正面显示和分析。因此,笔者提出了融合多尺度感知和残差连接的U型深度学习模型,并采用该模型实现了PACT光声图像中皮肤信号的智能分割。首先,提出以单类皮肤区域标注为基准标签图像的非像素级皮肤区域标注方法,该方法不需要像素级图像标注,能够显著降低数据处理的复杂度;然后,设计了皮肤完整性拟合和皮肤掩膜生成算法,并采用该算法实现了PACT图像中皮肤信号的自动去除。使用PACT成像实验获得的人体腿部外周血管光声图像验证了所提方法在皮肤组织高精度智能提取和去除方面的正确性和有效性。与现有的皮肤去除工作相比,本文所提皮肤去除算法对MAP图像的重建误差下降了50%~70%,峰值信噪比平均提升了约4.5 dB,为深层组织PACT图像的高清晰显示提供了一条有效途径。

关键词 生物光学; 光声成像; 皮肤分割; 深度学习; 外周血管成像

中图分类号 TP391 **文献标志码** A

DOI: 10.3788/CJL230930

1 引言

光声成像(PAI)是一种非侵入和非电离的生物医学成像方法,它结合了光学成像的高对比度和超声成像的深度优势,能够实现跨越细胞、组织和器官的多尺度、多分辨率成像^[1-3],在乳腺癌和甲状腺早期诊断、心血管易损斑块检测、脑功能研究等领域具有广阔的应用前景^[4-7]。光声计算层析成像(PACT)是一种重要的光声成像方式,与光声显微成像相比能够提供更深的组织穿透力。PACT通常使用超声阵列进行组织信号的探测,在配备多通道数据采集卡的情况下具有大视场下的高速成像潜力^[8]。因此,PACT可以实现小动物全身成像以及人体器官成像,具有巨大的临床应用潜力,目前在乳腺成像方面已经进入预临床应用研究阶段^[9-11]。

PACT已经展示了其在组织成像方面的大深度成像潜力。然而,皮肤组织中含有较多的黑色素,成像过程中皮肤会产生比较强的光声信号,遮挡了皮下感兴趣区域光声图像的正面高质量展示。因此,去除PACT图像中的皮肤信号,能够更清晰地观察深层组织的生理状态,为生命科学和疾病研究提供高质量的影像数据。

目前,几个科研小组已经开展了光声图像皮肤信

号去除的相关工作。早期,Maslov等^[12]对每一张二维剖面图像中的皮肤和血管进行了手动区分,但手动标注是一个非常耗时的过程,不适合高分辨率、较大区域下的大量B-Scan图像。因此,Zhang等^[13]基于光声显微成像系统,提出了一种自动检测皮肤轮廓的方法,即:对每一幅B-Scan图像,分析各个A-Line数据,选择相邻峰值差最大的信号作为皮肤位置,从而粗略地估计出皮肤的各个离散点;然后,通过非参数局部回归平滑方法拟合离散点,得到皮肤表面曲线。该方法能够有效去除光声显微图像中非常薄的表皮信号,但不适用于较厚的皮肤组织,而且在光照不均导致皮肤信号不易在背景中区分的情况下,会导致错误的皮肤信号提取。2020年,Baik等^[14]在自主开发的声学分辨率光声显微成像系统中,应用Zhang等的方法获得了高质量的人手掌和前臂的皮下血管网络结构图。近期,Wang等^[15]提出了一种基于Hessian矩阵的血管提取方法,而且他们认为在光声三维结构中,皮肤是点数据,是各向同性的,而血管是线结构,是各向异性的。因此,该方法的核心是将皮肤信号当作噪声点考虑,通过血管识别算法直接提取出血管结构。该方法本身不是皮肤去除算法,当PACT阵元相对稀疏时,图像中会存在显著的重建伪影,此时,易将伪影信号错误地提取为血管结构。

收稿日期: 2023-02-17; 修回日期: 2023-04-02; 录用日期: 2023-04-24; 网络首发日期: 2023-05-10

基金项目: 山东省自然科学基金(ZR2020MF105)、广东省生物医学光学重点实验室开放课题(2020B121201010)

通信作者: *jingmeng@qfnu.edu.cn; **cb.liu@sia.ac.cn

以上方法都是基于传统图像处理技术实现皮肤去除任务的。随着深度学习方法的快速发展,基于深度学习的图像分割成为当前的研究热点^[16-19]。目前,深度学习已被用于光声成像中,如稀疏采样光声图像重建^[20-21]、光声图像血管分割^[22-23]、光声图像去噪^[24-25]等。最近,Ly等^[26]将深度学习用于光声图像中的信号分类。在该工作中,Ly等基于光声显微系统对人类手和脚皮肤的微血管网进行了成像,并采用Unet深度学习模型在像素级别对微血管和皮肤进行了二分类。该工作实现了皮肤与皮肤区域内微血管的分离,但未开展深层组织成像中的皮肤信号去除研究。此外,像素级人工标注花费了大量时间。实验结果表明,该方法能够实现皮肤及其内部信号的有效分离,但皮肤和血管像素的误判率较高。

所查资料显示,目前已有工作能够去除光声图像中大部分皮肤信号的干扰,但它们大都基于浅层组织的光声显微图像进行皮肤去除或直接在背景中提取血管结构,而直接对深层组织PACT图像进行皮肤去除的相关工作还未见报道。笔者通过分析发现该工作存在以下2个难点:首先,由于PACT的非聚焦成像方式,采集到的光声信号需要通过重建算法才能得到光声图像,但图像中往往存在重建伪影,再加上PACT深层成像的单帧图像较大,导致当前像素级皮肤和血管分类的人工标注方法非常耗时;其次,由于重建伪影和光照强度变化的影响,皮肤区域的信号强度不均匀,存在诸多无法与背景进行区分的小段,增加了完整连续皮肤信号提取的难度。因此,笔者以PACT的深层组织高质量图像可视化为目标,探索智能皮肤提取的深度学习方法,直接基于二维图像处理实现单类皮肤信号的自动分割和去除。其中,针对第一个问题,考虑到皮肤的连续性和局部区域内皮肤厚度的一致性,提出了非像素级的皮肤区域标注方法,用于实现皮肤基准标签的快速生成;对第二个问题,笔者设计了融合残差和多尺度膨胀卷积的U型网络(MD-ResUnet),致力于提取更加完整连续的皮肤组织。为验证本文所提方法的有效性,开展了人体腿部的PACT光声成像实验,实验结果证明了所提出皮肤信号去除方法的正确性和高效性。

2 方 法

本文提出的PACT深层组织光声图像中皮肤信号智能去除的总体流程如图1所示。原始单帧图像归一化后和对应的基准标签图像作为网络输入,进行模型优化;然后,对预测得到的皮肤图像进行后处理和掩膜计算,从而重建得到无皮肤图像。整个过程包括三部分:1)基准标签图像的绘制;2)通过深度学习网络提取皮肤区域;3)掩膜生成与无皮肤图像重建。

在基准标签图像的绘制方面,鉴于像素级皮肤

和血管标注比较困难,考虑到皮肤组织的边界连续性和局部区域厚度的均一性,本文提出了单纯标记皮肤的方法,具体为:在基准标签图像中只标注皮肤组织,通过深度学习方法界定图像中的表皮区域。通过分析人体腿部的PACT光声图像,选择3个像素长度作为人工标注时的皮肤厚度,使用高级绘图软件中的钢笔绘图工具直接绘出皮肤区域。该方法与整幅图像的像素级标注相比显著降低了数据处理时间复杂度。

2.1 多尺度残差学习模型

根据PACT光声图像中皮肤信号的特点,本文提出了融合多尺度膨胀卷积和残差连接的U型深度学习模型(MD-ResUnet),该模型由编码器和解码器两部分组成,如图1所示。编码器和解码器部分的残差块(Res-block)依次进行了 3×3 卷积、批标准化(BN)和ReLU激活。卷积操作能够通过线性变换拟合局部信息;批标准化可将特征图的数值归一化到均值为0、标准差为1的分布,以改善梯度弥散,加快训练时模型的收敛速度;ReLU作为非线性函数可以发挥深度学习叠加层带来的优势,改善模型的拟合效果。模块中的残差连接将上层低级语义信息与提取的高级语义特征相加,补偿卷积过程中可能损失的特征信息。其中的 1×1 卷积用于配准卷积操作前后的特征图数量。在网络模型中,下层特征图数量是上一层的两倍,以便从上层传递下来的特征中提取更多的有用信息。此外,特征图尺寸的缩小间接增大了卷积时的感受野,可以使下层模块感知更大区域的语义信息。

为了拟合信号强度不同且不连续的皮肤表面,在编码器最底层设计了带残差的多尺度膨胀卷积块(Res-MD block),卷积核大小为 3×3 ,膨胀率分别设置为1、2和3。通过像素级的加和将三种膨胀率下提取的特征信息进行有效融合;残差结构的 1×1 卷积用于配准特征图数量,其与膨胀卷积后的结果加和避免了网络退化。本文提出的多尺度膨胀卷积块置于最底层的原因如下:在浅层模块中,由于特征图尺寸过大,需要较大的膨胀率才有可能探测到皮肤间断区域两端的信号,但大的卷积核会降低图像的分辨率;而最底层的模块具有最小尺寸的特征图,应用相对小的卷积模板就能覆盖皮肤间断区域,促进边界拟合。此外,皮肤边界存在着多处长度不一的难提取区域,即低对比度区域,这些区域最终会导致提取皮肤边界的断裂;而多尺度膨胀卷积能够感知不同视野下的信号特征,从而探测到不同大小弱皮肤区域两端的强皮肤信号,进而通过卷积操作提高皮肤提取的完整性和连续性。

2.2 掩膜生成与无皮肤图像重建

虽然通过深度学习模型已经获得了比较完整的皮肤区域图像,但仍然存在一些小的不连续片段,同时少量血管被预测为皮肤。为了得到完整连续的单纯皮

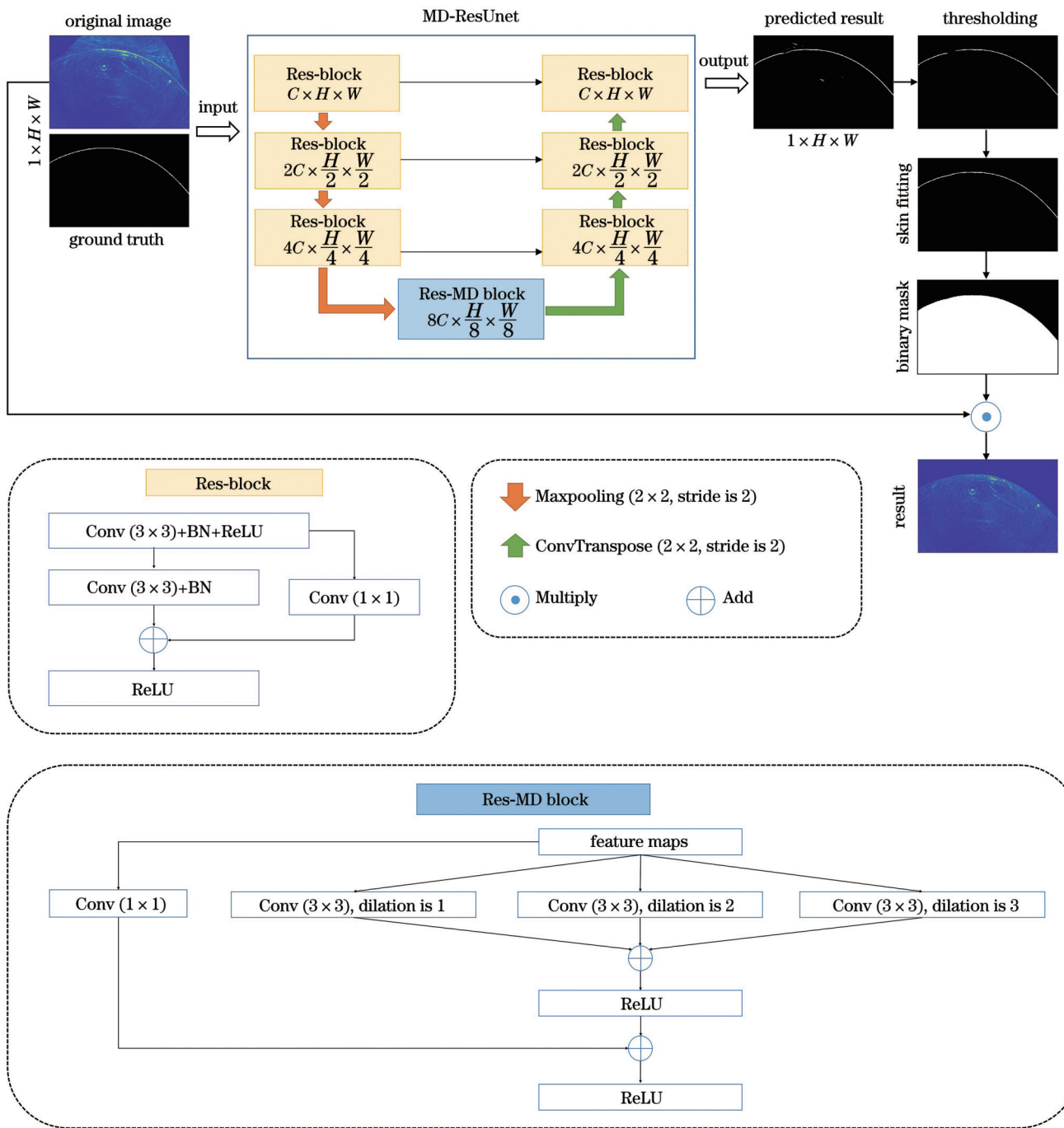


图1 基于深度学习的PACT皮肤信号去除方法

Fig. 1 PACT skin signal removal method based on deep learning

肤图像,笔者设计了相应的后处理操作,这些操作包括删除小连通区域和拟合不完整皮肤。

首先,用阈值方法将网络预测概率图转化为二值图像,本文选择阈值为0.5。然后,对于皮肤分割过程中出现的假阳性部分,采用去除小连通区域的方法进行过滤。具体地,如果某个值为1的像素点的8邻域仍然存在值为1的像素点,那么就认为这两个像素点在同一连通分量中。用这种方法可以得到每张B-Scan图像中的所有连通分量,进而可将小于一定阈值的小连通分量删除。本实验中设定小连通分量阈值为150 pixel。

2.2.1 皮肤完整性拟合

光声图像中的皮肤信号强度不均匀,信号较弱的皮肤区域不易在重建背景中分离,因此,部分皮肤分割结果存在弱信号区域皮肤断断续续的问题。针对这一问题,笔者设计了断点局部区域线性拟合的方法。根据断开的位置,将其分为皮肤中间断开和末端不完整提取两种,具体拟合方法如图2所示。如果断开的位置在皮肤中间,则断开间隙有两个端点,对应每个端点选取它的邻域皮肤信号(图中斜线填充方格),即像素值为1的4列数据,并根据邻域数据进行线性拟合。如果断开的位置在两端,就只将一个端点的邻域像素部

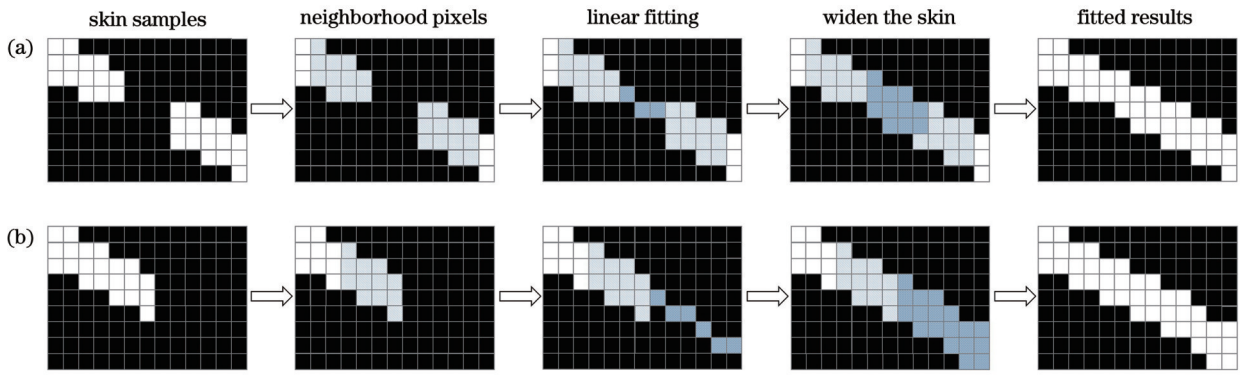


图 2 皮肤完整性拟合流程。(a)皮肤中间断开;(b)皮肤两端提取不完整

Fig. 2 Flowchart of skin-integrity fitting. (a) Breaking in the middle of skin; (b) incomplete skin at the ends

分进行拟合,将拟合后的直线连接到图像边缘;连接之后,补充的皮肤曲线为单像素(点状填充方格),然后采用将单像素点的上下像素点也赋值为1的方式增加皮肤厚度到3 pixel。经过上述操作,最终可以得到拟合后的完整连续的皮肤图像。在具体的实验中,将断裂皮肤两端邻域的皮肤像素作为一组已知数据点,使用Python软件SciPy库中的`curve_fit()`函数拟合断点。该函数基于最小二乘法通过最小化误差的平方和来拟合数据。

2.2.2 边界掩膜生成与皮肤去除

根据分割出的皮肤图像确定皮肤的下边界,然后生成皮肤去除掩膜。算法的具体流程如图3所示,共包括三个步骤:

1) 确定皮肤下边缘。对后处理后的二值皮肤图像自下而上逐列遍历,第一个为1的像素点确定为皮肤下边界上的点(点图案填充方格像素),遍历完成可得到皮肤下边缘的离散像素点。

2) 生成皮肤去除掩膜。根据获得的下边缘像素点,在二值皮肤图像中,对于每一列,令皮肤以上的像素值为0,以下的像素值为1(虚线边界方格像素及以下),得到二值掩膜图像。

3) 皮肤去除。将二值掩膜图像与原始采集的光声图像进行像素级点乘,重构出不受皮肤信号干扰的高对比度光声图像。

3 实验结果与分析

3.1 数据集及实验设置

本文使用的数据集为中国科学院深圳先进技术研究院提供的4组半环阵列-PACT系统采集的人体腿部外周血管光声数据,相关的系统说明和实验操作见文献[27]。每个数据集包含750帧,每帧大小为 $6\text{ cm} \times 4\text{ cm}$,成像区域的尺寸为 $15\text{ cm} \times 6\text{ cm}$ 。其中,50%的数据用于训练和验证网络,剩余的50%用于预测模型性能。基准标签图像由成像系统的专业科研人员进行标注。为了更好地捕捉皮肤表面的全局特征,模型训练过程中没有进行patch分割,而是将原始 $600\text{ pixel} \times 400\text{ pixel}$ 大小的B-Scan图像归一化后直接投入网络模型,以学习整体的皮肤轮廓。

本文中所有算法实验都是在Python 3.7中使用TensorFlow 2.2.0工具包实现的。软件安装在高性能计算机上,计算机配备16 GB NVIDIA TESLA P100 GPU。在网络模型优化过程中,批处理数量设置为8,采用Adam优化器对二分类交叉熵损失函数进行优化。学习率设置使用的是动态调整方法,即:初始学习率设置为 1×10^{-4} ,用较大步幅探寻极小值点,如果连续三轮验证集的损失函数值不下降,则将学习率更新为当前学习率的1/2,但最小不小于 1×10^{-6} ,以此逐渐向极小值点靠近。训练轮数设为100轮,保存验证集损失函数值最小的一轮模型为最优模型。

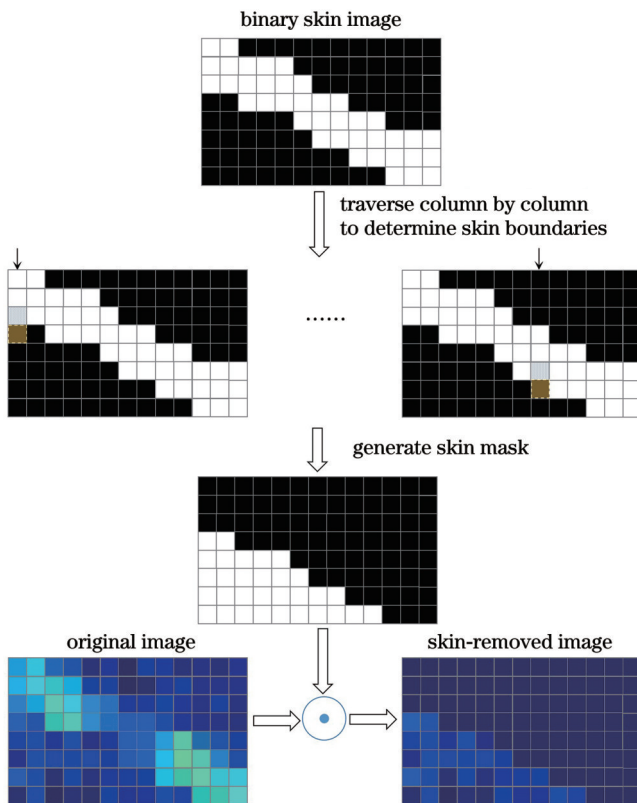


图 3 掩膜生成与皮肤去除流程图

Fig. 3 Flowchart of mask generation and skin removing

3.2 腿部外周血管成像

为验证所提深度学习方法的皮肤分割能力,开展了人体腿部外周血管光声图像皮肤提取实验。采集了 4 组光声图像数据,其中的两组用来优化网络模型,另两组用来预测网络性能,实验结果列在图 4 中。由于本文提出的 MD-ResUnet 网络是以 Unet 为基本骨架,添加了残差连接和多尺度膨胀卷积构成的,因此,在图 4 中展示了从骨架网络到本文所提网络结构的 B-Scan 图像的皮肤分割结果。前三行图像来自第一个数据集,后三行图像来自第二个数据集,在每个数据集中任意选取了三帧图像进行结果的可视化展示。第一列是原始的 B-Scan 图像,可以看到皮肤区域的信号比较强,在正面观察组织三维图像的情况下,皮肤干扰了皮下血管信号的清晰显示,如图 5(a)所示。第二列是

对应皮肤区域的基准标签图像。第三列到第五列分别对应 Unet^[28]、ResUnet^[29]和本文所提模型 MD-ResUnet 的皮肤图像分割结果。对比可以看出,Unet 网络模型能够提取出皮肤区域的大部分轮廓,但存在比较多的断开间隙,皮肤的连续性较差,如图 4(a3)~(f3)所示。将传统卷积操作改为残差连接后,皮肤区域的断开间隙明显减小,连续性得到了一定程度的提高,如图 4(a4)~(f4)所示。加入多尺度膨胀卷积块后,可以观察到大部分窄的皮肤分割间隙已经被拟合,而大的分割间隙有效缩短,如图 4(a5)~(f5)所示,提取的皮肤表面整体上更加光滑连续。对比可以发现,MD-ResUnet 在两组数据集的不同 B-Scan 图像上展示了良好的皮肤提取性能(如图中虚线框所示),证明了本文提出的深度学习模型在皮肤分割方面的可行性和有效性。

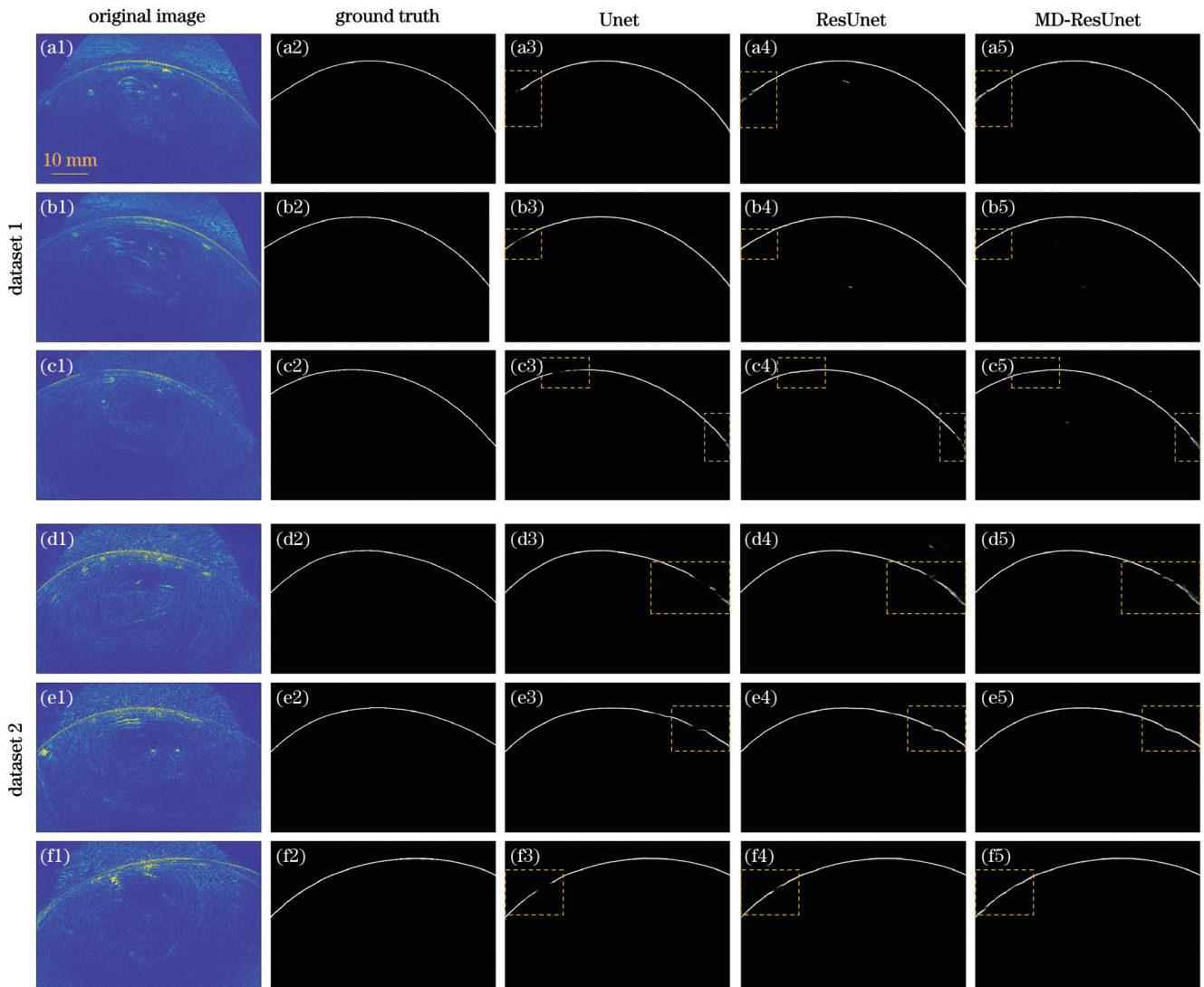


图 4 人体腿部外周血管 PACT 成像实验。(a1)~(f1)原始 PACT 图像;(a2)~(f2)基准标签图像;(a3)~(f3)Unet 模型的预测结果;(a4)~(f4) ResUnet 模型的预测结果;(a5)~(f5) MD-ResUnet 模型的预测结果

Fig. 4 PACT imaging for human periphery blood vessels of legs. (a1)~(f1) Original PACT images; (a2)~(f2) ground truth; (a3)~(f3) predicted results of Unet; (a4)~(f4) predicted results of ResUnet; (a5)~(f5) predicted results of MD-ResUnet

为了量化评估不同网络模型的皮肤分割性能,笔者对预测结果进行了多个参数的量化计算,这些参数

包括准确率(accuracy)、Dice 系数、灵敏度(sensitivity)、特异性(specificity)以及所有参数的平均值。从每个

数据集中选择 60 张 B-Scan 图像作为样本数据进行参数计算。由于连续帧之间图像数据的分布相似性较高,故而对于 60 个样本数据采用均匀间隔抽取的

方式,以保证参数计算的合理性和有效性。表 1 中列出了两个数据集下三种深度学习模型的评价指标值。

表 1 不同深度学习模型性能的量化评估

Table 1 Quantitative evaluation among different deep learning models

Dataset	Model	Accuracy / %	Dice coefficient / %	Specificity / %	Sensitivity / %	Average / %
Dataset 1	Unet	80.3924	99.6900	99.8733	87.6141	91.8925
	ResUnet	86.0536	99.7006	99.8640	88.9374	93.6389
	MD-ResUnet	86.7972	99.7073	99.8480	90.4299	94.1956
Dataset 2	Unet	69.7649	99.3782	99.7198	75.9882	86.2128
	ResUnet	72.4788	99.3696	99.6549	79.8348	87.8345
	MD-ResUnet	72.7585	99.3620	99.6283	81.1242	88.2183

准确率体现了预测正确的样本数量占总量的百分比,Dice 系数则体现了计算预测结果和标签之间的相似度。因此,从准确率和 Dice 系数可以看出三个模型的整体性能。在数据集 1 中,MD-ResUnet 的准确率比 Unet 和 ResUnet 分别高 6.4048 和 0.7436 个百分点;在数据集 2 中,MD-ResUnet 的准确率比 Unet 和 ResUnet 分别高 2.9936 和 0.2797 个百分点。在 Dice 系数方面,本文所提模型与 Unet、ResUnet 的差别不大,总体略高。

特异性和灵敏度这两个指标分别评价模型预测正样本和负样本的性能。在两个数据集中,Unet 的特异性整体比 ResUnet 和 MD-ResUnet 高,而 MD-ResUnet 的灵敏度整体比 Unet 和 ResUnet 高。这说明 Unet 具有较高的假阴性率,而 MD-ResUnet 具有更高的真阳性率。

最后,对各个参数进行平均值计算,综合评估不同算法的整体性能。在数据集 1 中,MD-ResUnet 模型比 Unet 高 2.3031 个百分点,比 ResUnet 高 0.5567 个百分点;在数据集 2 中,MD-ResUnet 模型比 Unet 高 2.0055 个百分点,比 ResUnet 高 0.3838 个百分点。这一结果从量化指标上展现本文模型的优越性。

3.3 对比实验

在传统方法中,比较经典的皮肤去除工作是 Zhang 等提出的局部加权拟合方法。本文方法与 Zhang 等的皮肤去除方法的对比结果如图 5 所示。图 5(a)和图 5(b)是两个数据集原始光声图像沿深度方向的最大振幅投影图像(MAP)。可以明显看出表皮信号对深层血管信号的遮盖,进而阻碍了感兴趣区域信息的三维正面显示。局部加权拟合方法对两个数据集的皮肤去除结果如图 5(c)和图 5(d)所示,可见该方法能够去除大部分的表皮信号,深层血管网络的清晰度显著提高。但观察图像可以发现一些区域仍然存在着表皮信号干扰,如矩形框区域所示。本文所提深度学习方法能够更加彻底

地去除皮肤信号,还原更加真实、清晰的深层组织结构,如图 5(e)和图 5(f)所示。为更加清楚地展示不同方法的效果,提取了一张典型 B-Scan 图像进行皮肤去除效果的对比,如图 5(a1)~5(f1)所示,并选取典型的局部区域进行放大显示。可以看出:Zhang 等的方法在某些 B-Scan 图像上的皮肤分割效果不理想,仍然存在大量断续的皮肤信号,如图 5(c1)和图 5(d1)所示;而本文提出的深度学习方法能够去除几乎全部的皮肤组织,这在图 5(e1)和图 5(f1)的放大子图中可以更直观地看出。以上对比实验展示了深度学习方法在皮肤分割工作中的优越性。

为了进一步分析不同算法下皮肤去除之后的图像质量,将算法计算得到的去皮肤 MAP 图与其对应的标签 MAP 图进行了均方误差(MSE)、结构相似性(SSIM)和峰值信噪比(PSNR)三方面的性能评估。均方误差可以反映标签 MAP 图与预测 MAP 图之间的差异程度,结构相似性可以反映两幅图像的相似程度,峰值信噪比是衡量图像质量的重要指标之一。不同数据集下的各指标值列在表 2 中。从表 2 中的数据可以看出,深度学习方法的均方误差指标比传统方法降低了约 50%~70%,峰值信噪比平均提高了约 4.5 dB,结构相似性也有明显提高。以上数据证明了去除皮肤后的 MAP 图能够有效去除皮肤干扰信号,显著提升图像的清晰度。

表 2 传统皮肤去除方法与所提深度学习方法的量化比较

Table 2 Quantitative comparison between traditional skin-removal method and proposed deep learning method

Dataset	Method	MSE	SSIM / %	PSNR / dB
Dataset 1	Zhang <i>et al.</i>	0.0136	98.6654	38.6774
	MD-ResUnet	0.0040	99.0471	43.9745
Dataset 2	Zhang <i>et al.</i>	0.0072	98.3557	38.4836
	MD-ResUnet	0.0032	99.2403	41.9694

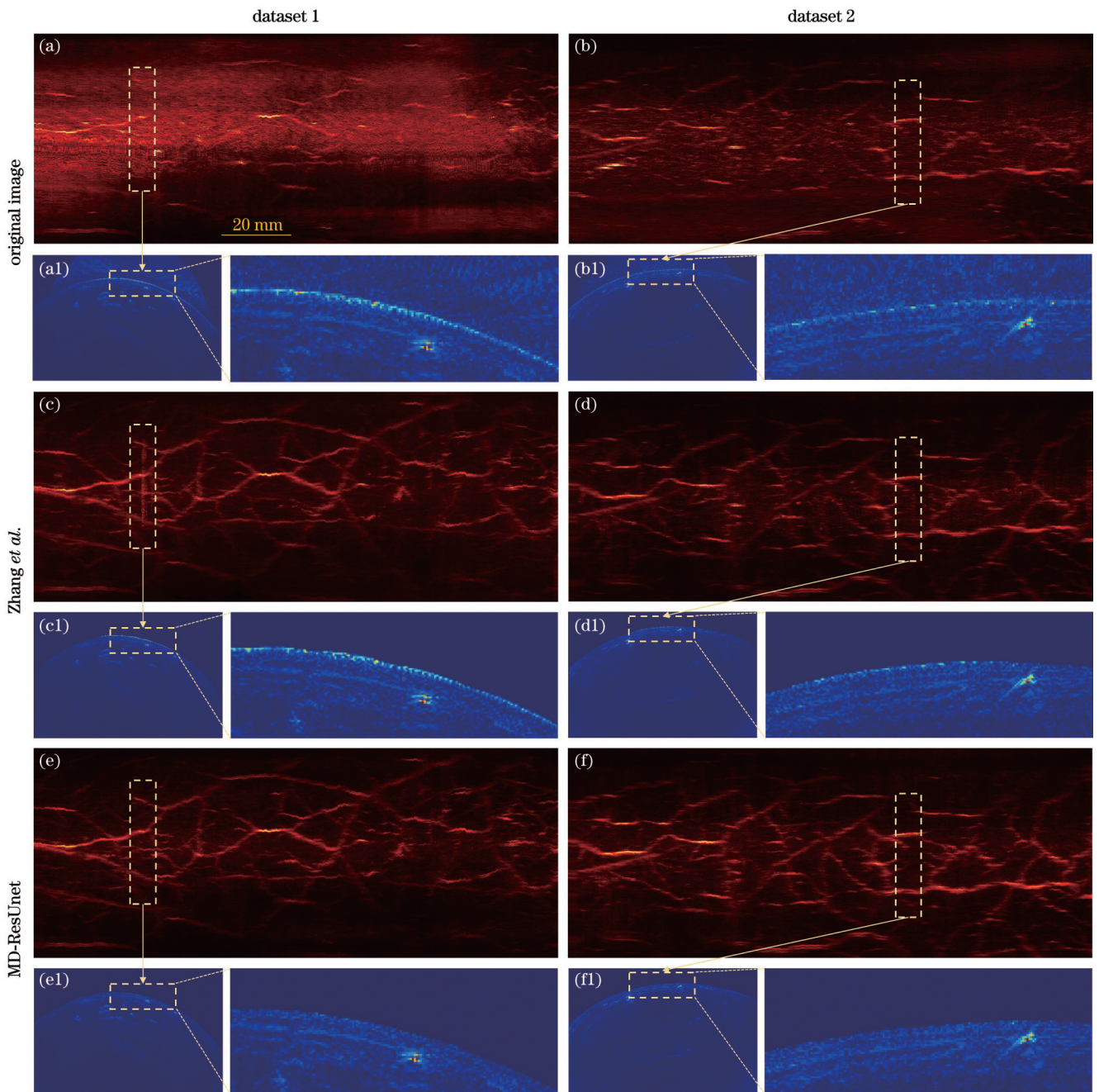


图 5 皮肤去除算法的比较。(a)(b)对应两个数据集的原始 PACT 的 MAP 图像;(c)(d)局域加权拟合方法的皮肤去除 MAP 图像;(e)(f) MD-ResUnet 方法的皮肤去除 MAP 图像;(a1)~(f1)典型的 B-Scan 图像及其放大子图

Fig. 5 Comparisons of different skin-removal methods. (a)(b) Original PACT MAP images from the two datasets; (c)(d) skin-removal MAP images with local weight fitting method; (e)(f) skin-removal MAP images with MD-ResUnet; (a1)~(f1) typical B-Scan images and their enlarged sub-images

4 结 论

本文基于 PACT 图像与人体皮肤特征,提出了一种以皮肤区域分割为核心的 PACT 深层组织图像皮肤去除方法,并设计了新型 U 型学习网络 MD-ResUnet 来实现皮肤分割任务。该方法以皮肤区域为基准标签图像,不需要像素级标注,降低了数据处理的复杂度。通过实验获得的人体腿部外周血管光声图像验证了本文所提皮肤去除方法的可行性和优越性,为

PACT 的深层组织高质量图像生成提供了一种有效方法。然而,本文提出的网络模型还无法实现完全连续的皮肤区域的提取,仍然存在着部分断开的皮肤间隙。对此,笔者将继续研究新型网络模型,致力于完全连续皮肤表面的生成。此外,本文实验是基于人体腿部外周血管进行成像的,该组织表面相对规则,皮肤区域整体形状是类弧形,便于深度学习掌握其整体形态和结构。对于一些复杂表面的成像组织,比如手指、手腕等,皮肤表面形状多变,同帧图像中存在着更显著的光

照差别,导致所成图像中皮肤区域的信号更加不均匀。笔者尝试着将本文所提算法用于人体手腕部 PACT 图像的皮肤去除工作,但由于手腕部位的皮肤不规则且凹凸不平,提取出的皮肤断裂区域较多。应用所提线性拟合方法能够得到完整的皮肤边界,实现皮肤去除功能;但由于拟合的断开区域较多,提取的皮肤精度降低。下一步,笔者将探索新的深度学习模型,以进一步提高皮肤提取的完整性。

参 考 文 献

- [1] 李有楠, 刘骧征, 王子滔, 等. 激光技术在血管内成像及治疗方面的应用[J]. 激光与光电子学进展, 2021, 58(23): 2300002.
Li Y N, Liu X Z, Wang Z T, et al. Application of laser technology in endovascular imaging and therapy[J]. Laser & Optoelectronics Progress, 2021, 58(23): 2300002.
- [2] 付武兵, 梁贻智, 仲晓轩, 等. 光纤光声血氧饱和度测量与功能成像[J]. 光学学报, 2022, 42(20): 2017001.
Fu W B, Liang Y Z, Zhong X X, et al. Optical fiber photoacoustic blood oxygen saturation measurement and functional imaging[J]. Acta Optica Sinica, 2022, 42(20): 2017001.
- [3] 赵鹏艳, 陈重江. 多模态光声成像技术及其在眼科学中的应用进展[J]. 激光与光电子学进展, 2022, 59(6): 0617014.
Zhao P Y, Chen Z J. Progress of multimodal photoacoustic imaging and its application in ophthalmology[J]. Laser & Optoelectronics Progress, 2022, 59(6): 0617014.
- [4] Dai Y F, Yu X, Wei J S, et al. Metastatic status of sentinel lymph nodes in breast cancer determined with photoacoustic microscopy via dual-targeting nanoparticles[J]. Light: Science & Applications, 2020, 9: 164.
- [5] Gui Y, Cheng K, Wang R J, et al. Photoacoustic detection of follicular thyroid carcinoma using targeted nano-Au-tripods[J]. Chinese Journal of Chemical Engineering, 2022, 44: 1-7.
- [6] Yang D Y, Zhu Y, Kong J Q, et al. "Light in and sound out": review of photoacoustic imaging in cardiovascular medicine[J]. IEEE Access, 2019, 7: 38890-38901.
- [7] Chen W T, Tao C, Hu Z Z, et al. Non-invasive and low-artifact *in vivo* brain imaging by using a scanning acoustic-photoacoustic dual mode microscopy[J]. Chinese Physics B, 2022, 31(4): 044304.
- [8] Wu Z G, Li L, Yang Y R, et al. A microrobotic system guided by photoacoustic computed tomography for targeted navigation in intestines *in vivo*[J]. Science Robotics, 2019, 4(32): eaax0613.
- [9] Xia J, Wang L V. Small-animal whole-body photoacoustic tomography: a review[J]. IEEE Transactions on Bio-Medical Engineering, 2014, 61(5): 1380-1389.
- [10] Lee C, Choi W, Kim J, et al. Three-dimensional clinical handheld photoacoustic/ultrasound scanner[J]. Photoacoustics, 2020, 18: 100173.
- [11] Kosik I, Brackstone M, Kornecki A, et al. Lipid-weighted intraoperative photoacoustic tomography of breast tumors: volumetric comparison to preoperative MRI[J]. Photoacoustics, 2020, 18: 100165.
- [12] Maslov K, Stoica G, Wang L V. *In vivo* dark-field reflection-mode photoacoustic microscopy[J]. Optics Letters, 2005, 30(6): 625-627.
- [13] Zhang H F, Maslov K I, Wang L V. Automatic algorithm for skin profile detection in photoacoustic microscopy[J]. Journal of Biomedical Optics, 2009, 14(2): 024050.
- [14] Baik J W, Kim J Y, Cho S, et al. Super wide-field photoacoustic microscopy of animals and humans *in vivo*[J]. IEEE Transactions on Medical Imaging, 2020, 39(4): 975-984.
- [15] Wang L, Ke D, Xin H S, et al. Optical-visualized photoacoustic tomographic navigation[J]. Applied Physics Letters, 2023, 122(2): 023701.
- [16] Chen C, Qin C, Qiu H Q, et al. Deep learning for cardiac image segmentation: a review[J]. Frontiers in Cardiovascular Medicine, 2020, 7: 25.
- [17] Chen C H, Chuah J H, Ali R, et al. Retinal vessel segmentation using deep learning: a review[J]. IEEE Access, 2021, 9: 111985-112004.
- [18] Ali S, Li J Q, Pei Y, et al. State-of-the-art challenges and perspectives in multi-organ cancer diagnosis via deep learning-based methods[J]. Cancers, 2021, 13(21): 5546.
- [19] Goni M R, Ruhaiyem N I R, Mustapha M, et al. Brain vessel segmentation using deep learning—a review[J]. IEEE Access, 2022, 10: 111322-111336.
- [20] 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.
- [21] Vu T, DiSpirito A, Li D W, et al. Deep image prior for undersampling high-speed photoacoustic microscopy[J]. Photoacoustics, 2021, 22: 100266.
- [22] Yuan A Y, Gao Y, Peng L L, et al. Hybrid deep learning network for vascular segmentation in photoacoustic imaging[J]. Biomedical Optics Express, 2020, 11(11): 6445-6457.
- [23] Zhao H X, Huang J A, Zhou Q A, et al. Deep learning-based optical-resolution photoacoustic microscopy for *in vivo* 3D microvasculature imaging and segmentation[J]. Advanced Intelligent Systems, 2022, 4(9): 2200004.
- [24] Allman D, Reiter A, Bell M A L. Photoacoustic source detection and reflection artifact removal enabled by deep learning[J]. IEEE Transactions on Medical Imaging, 2018, 37(6): 1464-1477.
- [25] Gulenko O, Yang H, Kim K, et al. Deep-learning-based algorithm for the removal of electromagnetic interference noise in photoacoustic endoscopic image processing[J]. Sensors, 2022, 22(10): 3961.
- [26] Ly C D, Nguyen V T, Vo T H, et al. Full-view *in vivo* skin and blood vessels profile segmentation in photoacoustic imaging based on deep learning[J]. Photoacoustics, 2022, 25: 100310.
- [27] Chen T, Liu L J, Ma X H, et al. Dedicated photoacoustic imaging instrument for human periphery blood vessels: a new paradigm for understanding the vascular health[J]. IEEE Transactions on Bio-Medical Engineering, 2022, 69(3): 1093-1100.
- [28] 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.
- [29] Zhang Z X, Liu Q J, Wang Y H. Road extraction by deep residual U-net[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(5): 749-753.

Intelligent Skin-Removal for Photoacoustic Computed Tomography

Yuan Yixin¹, Chen Tao², Liu Chengbo^{2**}, Meng Jing^{1*}

¹*School of Computer, Qufu Normal University, Rizhao 276826, Shandong, China;*

²*Institute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China*

Abstract

Objective Photoacoustic computed tomography (PACT) is an important photoacoustic imaging modality. Compared with photoacoustic microscopy, PACT can detect biological tissues located several centimeters deep without external contrast agents. Equipped with a multi-channel data acquisition card, PACT has the potential for high-speed imaging under a large field of view and is currently used in clinical and preclinical applications, such as whole-body imaging of small animals and human organs. However, skin tissue contains a lot of melanin, and the high-intensity photoacoustic signal from the skin covers the deep subcutaneous tissue information during the imaging process, hindering the *en-face* display and analysis of the photoacoustic image of the region of interest. Existing works have successfully removed most of the skin signals in photoacoustic images, but there are still some existing problems: (1) most of them are based on photoacoustic microscopic images of shallow tissues or directly extracted vascular structures in the background; the skin removal of deep tissue PACT images has not been reported; (2) the current pixel-level manual labeling takes a lot of time, and there are shortcomings of low extraction accuracy and low efficiency; (3) owing to reconstruction artifacts and changes in light intensity, the signal amplitudes of the skin area are uneven, and there exists many small segments that cannot be distinguished from the background, which increases the difficulty of extracting a complete and continuous skin signal.

Methods Considering the continuity of the skin tissue and the uniformity of the thickness of the local imaging area, this study proposes a U-shaped deep learning (DL) model that combines multi-scale perception and a residual structure (MD-ResUnet) to automatically remove skin areas in PACT deep tissue photoacoustic images. The introduction of the residual structure in this model can integrate low- and high-level feature information to prevent model degradation, and the multi-scale dilated convolution blocks can increase the continuity and integrity of skin extraction. In the skin segmentation task, a single-type skin region label was proposed as the ground truth, which significantly reduces the complexity of data annotation, compared with the previous pixel-level multi-type annotation. Subsequently, an algorithm of skin integrity fitting and skin mask generation was designed based on the extracted binary image of the skin, to realize the automatic removal of the skin signal in the PACT image. A total of four PACT datasets were used in our experiments, two of which were used for model optimization and two for experimental verification.

Results and Discussions The photoacoustic images of the peripheral blood vessels of human legs from PACT verified the correctness and effectiveness of the proposed method on high-precision extraction and removal of skin tissue. In the task of skin segmentation, the comparative experiments with the existing network models of Unet and Res-Unet, show that the DL model MD-ResUnet proposed in this study can fit most of the narrow skin segmentation gaps, effectively shorten the large segmentation gaps, and the extracted skin is overall more accurate, smooth and continuous (Fig. 4). Compared with the existing skin removal works, the deep learning method proposed in this study can thoroughly remove the skin signal and restore a more realistic and clear deep tissue structure (Fig. 5). Quantitative analysis shows that the reconstruction error of the skin-free image has dropped by 50%–70%, and the peak signal-to-noise ratio is averagely increased by 4.5 dB (Table 2), which may provide an effective method for the high-definition display of deep tissue PACT images.

Conclusions This study proposes a novel skin removal method for PACT deep tissue images with skin region segmentation as the core and designs a new U-shaped DL network MD-ResUnet to achieve the skin segmentation task. The proposed single-class skin-area labeling method significantly reduces the complexity of data processing, and the boundary fitting and mask generation methods realize the complete removal of skin areas, providing an effective method for high-quality deep tissue image generation in PACT. However, the network model proposed in this study cannot yet achieve fully continuous skin-region extraction, and there are still partially disconnected skin gaps. In addition, the experiment in this study is based on the imaging of the peripheral blood vessels of the human leg. The surface of this tissue is relatively regular, and the overall shape of the skin is arc-like, which is convenient for DL to grasp its overall structure features. For imaging tissues with complex surfaces, such as fingers and wrists, the surface shape of the skin is variable, and there will be more significant illumination differences in the same image frame, resulting in increased uneven skin area signals. In the future, we will explore advanced DL network models to implement the extraction of fully continuous skin surfaces in PACT images.

Key words bio-optics; photoacoustic imaging; skin segmentation; deep learning; imaging of periphery blood vessels