



抑制心血管图像序列中运动伪影的无监督深度学习方法

王茹, 孙正, 姚越

引用本文:

王茹, 孙正, 姚越. 抑制心血管图像序列中运动伪影的无监督深度学习方法[J]. *智能系统学报*, 2025, 20(4): 984-998.

WANG Ru, SUN Zheng, YAO Yue. Unsupervised deep learning method for suppressing motion artifacts in cardiac vascular image sequences[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(4): 984-998.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202408014>

您可能感兴趣的其他文章

基于智能计算的脑机制研究

Brain mechanism research based on intelligent computing

智能系统学报. 2021, 16(5): 850-856 <https://dx.doi.org/10.11992/tis.202103029>

基于反馈注意力机制和上下文融合的非模式实例分割

Feedback attention mechanism and context fusion based amodal instance segmentation

智能系统学报. 2021, 16(4): 801-810 <https://dx.doi.org/10.11992/tis.202007042>

利用残差密集网络的运动模糊复原方法

Image restoration with residual dense network

智能系统学报. 2021, 16(3): 442-448 <https://dx.doi.org/10.11992/tis.201912002>

基于深度学习的椎间孔狭窄自动多分级研究

Deep learning based automatic multi-classification algorithm for intervertebral foraminal stenosis

智能系统学报. 2019, 14(4): 708-715 <https://dx.doi.org/10.11992/tis.201806015>

高斯核函数卷积神经网络跟踪算法

Convolutional neural network tracking algorithm accelerated by Gaussian kernel function

智能系统学报. 2018, 13(3): 388-394 <https://dx.doi.org/10.11992/tis.201612040>

基于医学征象和卷积神经网络的肺结节CT图像哈希检索

Hashing retrieval for CT images of pulmonary nodules based on medical signs and convolutional neural networks

智能系统学报. 2017, 12(6): 857-864 <https://dx.doi.org/10.11992/tis.201706035>

DOI: 10.11992/tis.202408014

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20250224.1852.014>

抑制心血管图像序列中运动伪影的 无监督深度学习方法

王茹¹, 孙正^{1,2}, 姚越¹

(1. 华北电力大学 电子与通信工程系, 河北 保定 071003; 2. 华北电力大学 河北省电力物联网技术重点实验室, 河北 保定 071003)

摘要: 血管内超声 (intravascular ultrasound, IVUS) 和光学相干断层成像 (optical coherence tomography, OCT) 是诊断冠状动脉粥样硬化性病变的重要手段, 但心脏运动和血流搏动会产生运动伪影, 影响图像质量。为解决这一问题, 本文提出一种无监督深度学习方法, 用于抑制 IVUS/OCT 图像序列中的运动伪影。设计一个深度神经网络, 包含特征提取、上采样、运动估计和运动校正模块, 实现从连续回撤导管采集的图像序列到去伪影图像序列的映射。利用临床 IVUS/OCT 图像进行无监督训练, 实验结果表明, 该方法能显著提高纵向视图中管壁边缘的平滑度, 标准差和帧间差异度指标得到显著改善。与其他方法相比, 本文方法可使平均帧间差异度降低约 23%, 标准差降低约 24%。该方法有效解决了 IVUS/OCT 图像序列因运动伪影造成的质量下降问题, 优化了图像视觉效果。

关键词: 心血管成像; 冠状动脉内成像; 图像序列; 运动伪影; 运动补偿; 神经网络; 无监督学习; 弹性配准

中图分类号: TP391.41; R445 **文献标志码:** A **文章编号:** 1673-4785(2025)04-0984-15

中文引用格式: 王茹, 孙正, 姚越. 抑制心血管图像序列中运动伪影的无监督深度学习方法 [J]. 智能系统学报, 2025, 20(4): 984-998.

英文引用格式: WANG Ru, SUN Zheng, YAO Yue. Unsupervised deep learning method for suppressing motion artifacts in cardiac vascular image sequences[J]. CAAI transactions on intelligent systems, 2025, 20(4): 984-998.

Unsupervised deep learning method for suppressing motion artifacts in cardiac vascular image sequences

WANG Ru¹, SUN Zheng^{1,2}, YAO Yue¹

(1. Department of Electronic and Communication Engineering, North China Electric Power University, Baoding 071003, China;

2. Hebei Key Laboratory of Power Internet of Things Technology, North China Electric Power University, Baoding 071003, China)

Abstract: Intravascular ultrasound (IVUS) and optical coherence tomography (OCT) are important tools for diagnosing coronary atherosclerotic lesions. However, cardiac motion and pulsatile blood flow can cause motion artifacts that affect image quality. To address this issue, an unsupervised deep learning method for suppressing motion artifacts in IVUS/OCT image sequences is proposed in this paper. A deep neural network consisting of feature extraction, up-sampling, motion estimation, and motion correction modules is designed, and it can map continuous pullback image sequences to artifact-free image sequences. The network is trained in an unsupervised manner using clinical IVUS/OCT images. Experimental results demonstrate that this method remarkably improves the smoothness of the vessel wall boundaries in the longitudinal view, with substantial improvements in standard deviation and interframe dissimilarity metrics. Compared with other methods, the proposed approach reduces average interframe dissimilarity and standard deviation by approximately 23% and 24%, respectively. This method effectively solves the image quality degradation caused by motion artifacts in IVUS/OCT image sequences and optimizes the visual quality of the images.

Keywords: cardiac vascular imaging; intracoronary imaging; image sequences; motion artifact; motion compensation; neural network; unsupervised learning; elastic registration

收稿日期: 2024-08-15. 网络出版日期: 2025-02-25.

基金项目: 国家自然科学基金项目 (62071181).

通信作者: 孙正. E-mail: sunzheng@ncepu.edu.cn.

血管内超声 (intravascular ultrasound, IVUS) 和血管内光学相干断层成像 (intravascular optical coherence tomography, IVOCT) 是临床诊治冠状动脉

粥样硬化性病变的关键影像手段^[1-3]。二者成像原理类似,都是在X射线透视指导下,将导丝插入靶血管腔内并固定在远端,然后沿导丝将顶端带有探头的特制导管插入管腔中并推送至远端。在从远端向近端缓慢回撤导管的过程中,探头向周围组织发射超声波或低相干宽带红外激光,探测器采集周围组织反射/散射的超声波或光信号,再根据信号特性重建图像,完整显示血管横断面。

运动伪影是冠状动脉内成像临床应用的主要挑战。心脏的周期性收缩与舒张以及管腔内的搏动血流可能引起导管顶端相对于管腔的侧向位移,导致B模式图像中的血管结构发生变形和相邻帧之间血管横截面的错位,以及沿血管长轴方向的纵向视图中管壁呈现锯齿形。运动伪影不仅严重影响图像的视觉效果,还会导致关键信息的丢失,进而影响血管形态的定量测量^[4]、血流动力学分析^[5]、组织成分的定量分析^[6]以及三维重建^[7]的精度。尽管最新的频域IVOCT技术能将图像采集速度提升至100 f/s以上^[8],但在对整段血管进行成像时,导管与血管壁之间的相对运动仍会影响成像精度。因此,抑制冠状动脉内图像序列中的运动伪影,对于改善图像的视觉效果和为冠心病的计算机辅助诊断提供准确信息具有重要作用。

门控法是临床心脏成像中抑制运动伪影的主要技术,分为前瞻性门控和回顾性门控两种方式。前瞻性门控利用心电图(electrocardiogram, ECG)触发装置,在心脏的特定相位(如R波)采集信号或图像,形成门控图像序列^[9-10]。该方法需要特殊的触发采集装置,且由于每个心动周期只采集一帧图像,会显著延长介入操作的时间。回顾性门控则通过硬件和软件两种方式实现。硬件门控是在连续回撤导管采集图像序列的同时同步记录ECG信号,并在采集结束后,对照ECG信号选择相同心脏相位的图像组成门控帧序列^[11]。这种方法不会延长介入检查时间,但是ECG信号和介入图像的采集可能不完全同步,影响门控的准确性。软件门控则运用信号处理和图像分析技术,从连续采集的图像序列中提取心脏时相信息,进而选择出相同相位的图像组成门控序列。这种方法克服了ECG门控的缺点,适用于没有同步记录ECG信号的图像序列^[12]。例如,采用近邻传播(affinity propagation, AP)聚类算法对连续帧图像之间的相关矩阵进行聚类,提取出静态帧^[13]。

近年来,深度学习在心脏图像序列后处理中显示出巨大潜力,为回顾性门控提供了新的解决

方案^[14-15]。例如,基于无监督学习的IBG-Net(image-based gating network)采用临床采集的图像数据集对网络模型进行无监督训练,使其从连续回撤导管采集的原始图像序列中自动提取门控帧^[15]。然而,门控技术仅保留特定心脏相位的图像,可能丢失关键的诊疗信息,因此限制了其在动脉生物力学特性和血管壁弹性研究中的应用。

考虑到门控技术的缺点,直接抑制方法成为一种替代方案。这种方法不仅可以保证图像数据集的完整性,还能提供更全面的心脏成像信息。运动跟踪和图像配准是两种主要的解决方案。在运动跟踪领域,经典方法如光流法^[16]和块匹配法^[17]已被成功应用于IVUS图像序列的运动补偿。数据驱动的光流估计和块匹配方法也在心脏成像领域取得了显著进展。例如,基于监督学习的PWC-Net(CNNs for optical flow using pyramid, warping, and cost volume)^[18]和无监督的UPFlow(upsampling pyramid for unsupervised optical flow learning)^[19]有效提高了光流估计的准确性和鲁棒性。但是,光流法对超声图像中的血液随机运动和噪声非常敏感^[20],块匹配法的精度则受限于相似度计算的准确性,由于通常根据像素亮度计算相似度,因此其准确度受到反射信号干涉模式的帧间变化或者超声散斑的影响。此外,由于需要对参数空间进行全面搜索,计算成本也较高。

图像配准法通过分析图像特征或血管结构的周期性变化^[21]来提取运动参数,目的是消除相邻帧之间的运动差异。经典方法包括特征匹配和优化方法,它们通过检测图像中的特征点^[22]或计算相似性度量^[21,23]实现图像配准。经典方法在实现运动图像的精确对齐方面具有重要作用,但对噪声和图像变形较为敏感,在图像质量较差或存在大量噪声的情况下可能降低配准的准确性。基于深度学习的运动图像配准方法包括监督学习和无监督学习两类。监督学习方法如DeepReg(deep learning toolkit for medical image registration)^[24]、ST-FCN(spatial transformer fully connected network)^[25]和FC-DenseReg(fully convolutional dense shape regression)^[26],通过用大量带有配对标签的数据训练神经网络^[27-28],使其学习图像之间的变换关系,能够实现较高的配准精度。无监督学习方法则是利用神经网络模型学习图像之间的变换关系,通过最小化图像之间的差异或引入特定配准损失函数训练网络,无需带标签的配对样本数据^[29-31]。例如基于生成对抗网络(generative adversarial networks, GAN)架构的MAC-Net(motion artifacts correction net)^[32],通过计算图像之间的变形场抑制运

动伪影。MAC-Net 是采用仿真数据集训练的,而仿真数据与在真实场景中采集的图像数据之间存在差距,可能导致网络在处理植入支架后的血管段、严重硬化导致的血管弹性丧失和分叉病变等复杂场景时面临挑战。

本文提出一种直接抑制冠状动脉内图像序列中运动伪影的无监督深度学习方法,实现从连续回撤导管采集的图像序列到抑制伪影后图像序列的直接映射。采用临床采集的 IVUS/OCT 图像序列对算法进行无监督训练,并与基于深度学习的离线门控法、基于图像配准的非学习方法、基于近邻传播聚类的离线门控法以及 MAC-Net 进行比较,以验证其在抑制运动伪影方面的优越性。

1 抑制冠状动脉内图像序列中运动伪影的深度学习方法

本文设计并构建一个深度神经网络,命名为弹性图像配准网络 (elastic image registration network, EIR-Net)。EIR-Net 通过估计图像序列中相邻帧 B 模式图像中对应像素之间的运动场,实现二者之间的弹性配准,从而有效抑制运动伪影。

如图 1 所示,对于连续采集的图像序列,从首帧 (第 $i=1$ 帧) 开始,将当前帧 (第 i 帧) 作为校正目标图像,后一帧 (第 $i+1$ 帧) 作为待校正图像,依次输入训练后的 EIR-Net 中。网络输出校正后的第 $i+1$ 帧图像。通过重复该过程,最终得到抑制运动伪影之后的图像序列。

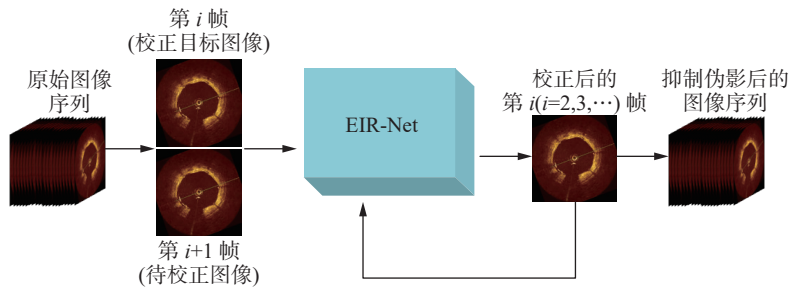


图 1 采用无监督学习框架 EIR-Net 对冠状动脉内图像序列进行运动伪影抑制的方法流程

Fig. 1 Method flow of motion artifact suppression in intracoronary image sequences using the unsupervised learning framework EIR-Net

1.1 EIR-Net 的架构

如图 2 所示, EIR-Net 由 4 个关键部分组成: 特征提取模块、运动估计模块、上采样模块和运动校正模块。

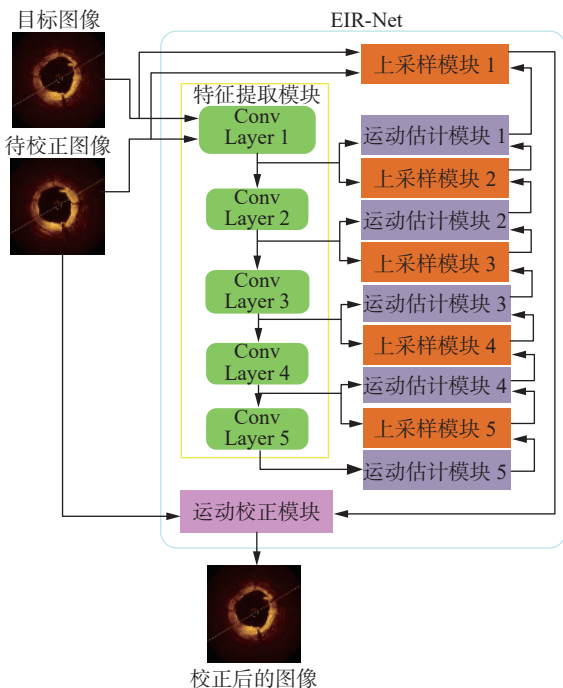


图 2 EIR-Net 的结构

Fig. 2 Architecture of EIR-Net

首先,输入图像通过特征提取模块来提取多尺度的特征。接着,运动估计模块根据特征提取模块输出的特征图估计低分辨率运动场,在随后的每一层中,该模块会基于前一层上采样后的运动场并结合该层的特征信息,对运动场进行细化。然后,上采样模块将低分辨率的运动场逐层放大,并与高分辨率的图像特征结合,在逐步提高分辨率的过程中,细化运动场,从而生成更高分辨率的运动场。最终,通过运动校正模块将运动场作用于待校正图像上,实现运动场到校正后图像的转换。

1.1.1 特征提取模块

特征提取模块包括 5 个层级,旨在分别从两幅输入图像中提取不同尺度的特征对。如图 3 所示,该模块由一对结构相同的孪生网络构成,分别用于从输入的待校正图像 I_m 和目标图像 I_t 中提取出 5 组不同尺度的特征对,即 $(F_{m,l}, F_{t,l})$, 其中 $l=1, 2, 3, 4, 5$ 。通过多个层级,可以获取不同尺度上的特征对,较低层级的特征图具有较大的感受野,能够捕捉全局信息,而较高层级的特征图则具有较小的感受野,有助于捕捉局部细节。然而,层级数过多会增加计算复杂度,提升训练

成本, 并增加过拟合的风险; 层级数过少则可能导致信息丢失, 无法充分提取不同尺度上的特征, 从而影响模型的性能和泛化能力。因此, 在实际应用中, 考虑到精度和效率, 层级数通常选择 5^[18-19,33]。

每个孪生网络包含 5 个卷积层和 5 个下采样

层。卷积采用 $3 \times 3 \times T$ 的滤波核, 步长为 1。下采样使用 $3 \times 3 \times T$ 的滤波核, 步长为 2, 填充为 same。其中, T 表示每层卷积核的数量, 依次为 16、32、64、96 和 128。每个卷积层之后都应用带泄漏修正的线性整流函数 (leaky rectified linear unit, Leaky ReLU) 作为激活函数。

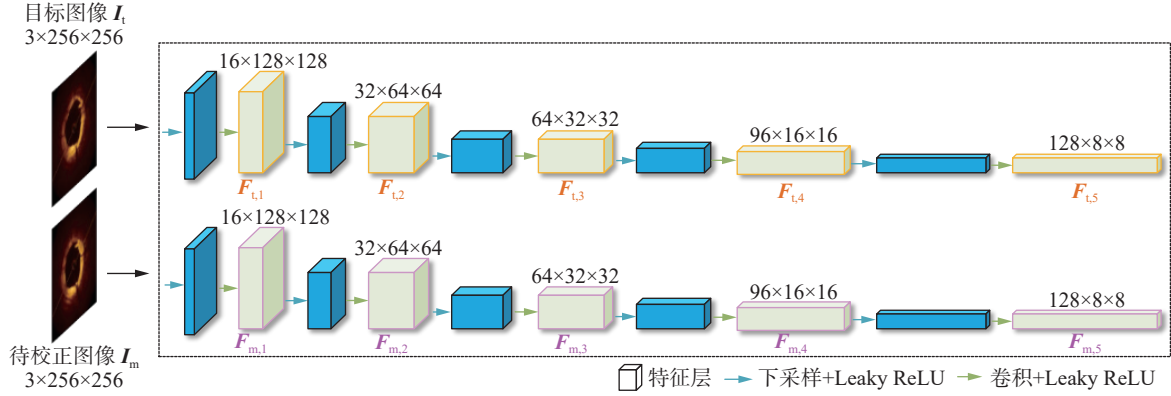


图 3 特征提取模块的结构

Fig. 3 Structure of the feature extraction module

1.1.2 运动估计模块

与特征提取模块的 5 层结构相对应, EIR-Net 中共包含 5 个运动估计模块。其中第 1 个模块的构成如图 4(a) 所示, 第 2~5 个模块的构成如

图 4(b) 所示, 第 1 个模块相比其他模块增加了上下文网络, 能够利用周围像素细化光流场, 提高运动估计的精度。运动估计模块用于根据特征提取模块输出的特征图估计低分辨率运动场。

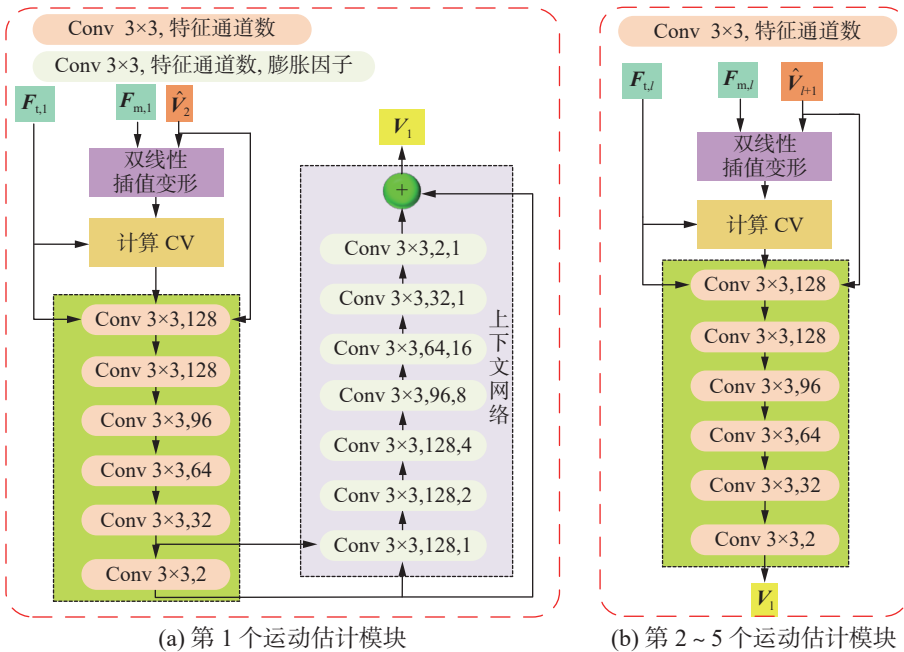


图 4 运动估计模块的结构

Fig. 4 Structure of motion estimation modules

第 $l(l=1, 2, 3, 4)$ 个运动估计模块的输入包括特征提取模块中第 l 层提取的输入图像的特征图对 $(F_{m,l}, F_{t,l})$ 以及第 $l+1$ 个上采样模块输出的上采样运动场 \hat{V}_{l+1} 。具体操作为: 首先, 利用 \hat{V}_{l+1} 对

$F_{m,l}$ 进行双线性插值变形, 得到

$$\bar{F}_{m,l} = F_{m,l}(x + \hat{V}_{l+1}(x)) = \sum_{x' \in N(x + \hat{V}_{l+1}(x))} F_{m,l}(x') \prod_{d \in \{0,1\}} [1 - |x_d + \hat{V}_{l+1,d}(x) - x'_d|]$$

式中: \mathbf{x} 是 $\mathbf{F}_{m,l}$ 中的像素, $\bar{\mathbf{F}}_{m,l}$ 是 \mathbf{x} 处经过双线性插值变形后的特征图, $\hat{\mathbf{V}}_{l+1}$ 是第 $l+1$ 个上采样模块输出的上采样运动场 (由水平分量和垂直分量组成), $\mathcal{N}(\mathbf{x} + \hat{\mathbf{V}}_{l+1}(\mathbf{x}))$ 是像素 $\mathbf{x} + \hat{\mathbf{V}}_{l+1}(\mathbf{x})$ 的 4-邻域, \mathbf{x}' 表示该邻域中的一个像素, $\mathbf{F}_{m,l}(\mathbf{x}')$ 是 \mathbf{x}' 的像素值, $d \in \{0, 1\}$ 是运动向量两个分量 (水平和垂直) 的标号, \mathbf{x}_d 和 \mathbf{x}'_d 分别表示 \mathbf{x} 和 \mathbf{x}' 的第 d 个分量。该式通过对邻域像素值的加权求和估算新像素值。 $1 - |\mathbf{x}_d + \hat{\mathbf{V}}_{l+1,d}(\mathbf{x}) - \mathbf{x}'_d|$ 是 $\hat{\mathbf{V}}_{l+1,d}(\mathbf{x})$ 的线性函数, 权重项 $\prod_{d \in \{0,1\}} [1 - |\mathbf{x}_d + \hat{\mathbf{V}}_{l+1,d}(\mathbf{x}) - \mathbf{x}'_d|]$ 是多个线性函数的乘积, 因此该项是可微的。最后, 求和部分是对各邻域点的线性组合, 因此整体表达式是可微的, 确保了网络在训练过程中可以通过反向传播优化运动场 $\hat{\mathbf{V}}_{l+1}$ 。

计算 $\mathbf{F}_{t,l}$ 与 $\bar{\mathbf{F}}_{m,l}$ 之间的相关度, 用代价体函数 (cost volume, CV) 表示:

$$C_{V,l}(\mathbf{x}_t, \mathbf{x}_m) = \frac{1}{N} (\mathbf{F}_{t,l}(\mathbf{x}_t))^T \bar{\mathbf{F}}_{m,l}(\mathbf{x}_m) \quad (1)$$

式中: \mathbf{x}_t 和 \mathbf{x}_m 是 $\mathbf{F}_{t,l}$ 和 $\bar{\mathbf{F}}_{m,l}$ 中的像素位置, N 是列向量 $\mathbf{F}_{t,l}(\mathbf{x}_t)$ 的长度。式 (1) 中的 \mathbf{x}_t 和 \mathbf{x}_m 满足条件:

$$|\mathbf{x}_t - \mathbf{x}_m|_{\infty} \leq C_V$$

式中: $|\mathbf{x}_t - \mathbf{x}_m|_{\infty}$ 是 \mathbf{x}_t 和 \mathbf{x}_m 之间的无穷范数距离, C_V 是计算 CV 的有限范围, 从第 1 层到第 5 层, C_V 值依次为 16、8、4、2 和 1。

然后, 采用 6 个级联的 3×3 卷积层 (每层卷积核的数量依次为 128、128、96、64、32 和 2, 除最后一层外, 每个卷积层之后均接有 Leaky ReLU 激活函数) 对 $\mathbf{F}_{t,l}$ 、 $C_{V,l}$ 和 $\hat{\mathbf{V}}_{l+1}$ 进行多尺度特征提取^[17], 生成粗运动场 \mathbf{V}_l 。对于第 2、3、4 个运动估计模块, $\mathbf{V}_l (l=2, 3, 4)$ 即为该模块输出的低分辨率运动场。对于第 1 个运动估计模块, 粗运动场和倒数第 2 个卷积层 (Conv 3×3 , 32) 输出的特征图一并输入上下文网络, 最终输出低分辨率运动场 \mathbf{V}_1 。如图 2 所示, 第 5 个运动估计模块的输入仅包括 $(\mathbf{F}_{m,5}, \mathbf{F}_{t,5})$, 具体操作是: 首先利用式 (1) 计算 $\mathbf{F}_{t,5}$ 与 $\mathbf{F}_{m,5}$ 之间的相关度 $C_{V,5}$; 然后将 $\mathbf{F}_{t,5}$ 和 $C_{V,5}$ 一起输入级联卷积层中进行多尺度特征提取, 最终输出低分辨率运动场 \mathbf{V}_5 。

上下文网络是一个前馈卷积神经网络, 通过结合粗运动场和特征图, 利用周围像素对运动场进行修正和精细化, 从而提高运动估计的准确性。该网络由 7 个级联的 3×3 卷积层构成^[18], 卷积核步长为 1, 每层卷积核的数量依次为 128、128、128、96、64、32 和 2, 膨胀因子依次为 1、2、4、8、16、1 和 1。除最后一层外, 每层卷积之后均

接有 Leaky ReLU 激活函数。

1.1.3 上采样模块

EIR-Net 中共包含 5 个上采样模块。如图 5 所示, 每个上采样模块由加权变形、双线性插值变形、加权组合和密集块组成^[19]。

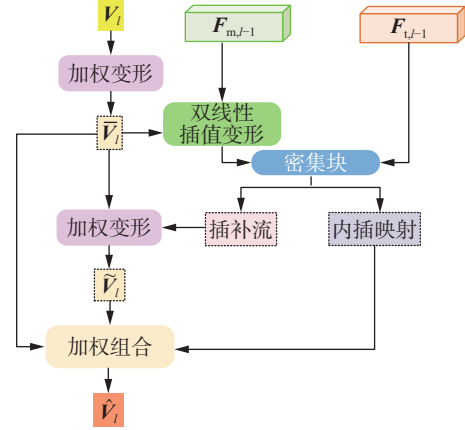


图 5 上采样模块的结构^[28]

Fig. 5 Structure of the upsampling module^[28]

第 1 个上采样模块的输入包括待校正图像 \mathbf{I}_m 、目标图像 \mathbf{I}_t (即底层特征 $(\mathbf{F}_{m,0}, \mathbf{F}_{t,0})$), 以及第 1 个运动估计模块输出的低分辨率运动场 \mathbf{V}_1 。第 $l (l=2, 3, 4, 5)$ 个上采样模块的输入则是 $\mathbf{F}_{m,l-1}$ 、 $\mathbf{F}_{t,l-1}$ 以及第 l 个运动估计模块输出的低分辨率运动场 \mathbf{V}_l , 其输出为上采样运动场 $\hat{\mathbf{V}}_l$ 。

第 l 个上采样模块的具体操作如下, 首先, 对 \mathbf{V}_l 进行加权变形, 生成较高分辨率的初始上采样运动场:

$$\bar{\mathbf{V}}_l(\mathbf{p}) = \sum_{\mathbf{k} \in \mathcal{N}(\mathbf{q})} w(\mathbf{q}, \mathbf{k}) \cdot \mathbf{V}_l(\mathbf{k})$$

其中

$$\mathbf{q} = \mathbf{p}/s$$

式中: \mathbf{p} 是较高分辨率运动场中的像素位置, s 是比例放大倍数, $\mathcal{N}(\mathbf{q})$ 是由与 \mathbf{q} 相邻的最近的 4 个像素组成的集合 (即 4-邻域), $w(\mathbf{q}, \mathbf{k})$ 是对 \mathbf{V}_l 进行双线性插值的权重, \mathbf{k} 是集合 $\mathcal{N}(\mathbf{q})$ 中的一个像素。

接着, 利用 $\bar{\mathbf{V}}_l$ 对 $\mathbf{F}_{m,l-1}$ 进行双线性插值变形, 得到

$$\bar{\mathbf{F}}_{m,l-1} = \mathbf{F}_{m,l-1}(\mathbf{p} + \bar{\mathbf{V}}_l(\mathbf{p})) = \sum_{\mathbf{p}' \in \mathcal{N}(\mathbf{p} + \bar{\mathbf{V}}_l(\mathbf{p}))} \mathbf{F}_{m,l-1}(\mathbf{p}') \prod_{d \in \{0,1\}} [1 - |\mathbf{p}_d + \bar{\mathbf{V}}_{l,d}(\mathbf{p}) - \mathbf{p}'_d|]$$

式中: $\bar{\mathbf{F}}_{m,l-1}$ 是像素 \mathbf{p} 处经过双线性插值变形后的特征图, $\bar{\mathbf{V}}_l(\mathbf{p})$ 是像素 \mathbf{p} 处的初始上采样运动场 (由水平分量和垂直分量组成), $\mathcal{N}(\mathbf{p} + \bar{\mathbf{V}}_l(\mathbf{p}))$ 是像素 $\mathbf{p} + \bar{\mathbf{V}}_l(\mathbf{p})$ 的 4-邻域, \mathbf{p}' 是该邻域中的像素, \mathbf{p}_d 和 \mathbf{p}'_d 分别是 \mathbf{p} 和 \mathbf{p}' 的第 d 个分量。

然后, 将变形后的特征图 $\bar{\mathbf{F}}_{m,l-1}$ 与目标图像的

特征图 F_{i-1} 一起输入密集块。密集块输出一个三通道张量图, 其中前两个通道表示插补流 U_i , 最后一个通道通过 sigmoid 层形成内插映射 B_i , 用于确定权重。接着, 用插补流 U_i 对 \bar{V}_i 进行加权变形, 得到

$$\tilde{V}_i(\mathbf{p}) = \sum_{\mathbf{k} \in N(\mathbf{q})} w(\mathbf{q}, \mathbf{k}) \cdot \bar{V}_i(\mathbf{k})$$

其中

$$\mathbf{q} = \mathbf{p} + U_i(\mathbf{p})$$

式中: $\tilde{V}_i(\mathbf{p})$ 是 \bar{V}_i 经过插补流加权变形之后的结果, $w(\mathbf{q}, \mathbf{k})$ 是权重。

最后, 对 \bar{V}_i 和 \tilde{V}_i 进行加权组合生成最终的上采样运动场:

$$\hat{V}_i = B_i \odot \bar{V}_i + (1 - B_i) \odot \tilde{V}_i$$

式中“ \odot ”是 Hadamard 乘积。由于双线性插值模糊主要发生在目标边缘区域, 因此内插映射 B_i 确保模型在这些边缘区域进行有效的插值流学习。

1.1.4 运动校正模块

如图 6 所示, 运动校正模块由网格生成器和采样器两部分组成。其输入包括待校正图像以及校正目标图像与待校正图像之间的运动场, 输出为校正后的图像。

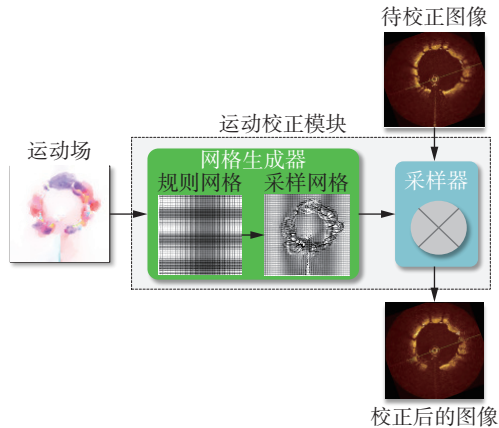


图 6 运动校正模块的结构

Fig. 6 Structure of the motion correction module

具体操作过程为: 首先, 网格生成器根据输入待校正图像的尺寸生成规则网格, 在运动场中查找每个网格点对应的位移向量, 并根据该位移向量将每个网格点调整到新的位置, 得到采样网格。接着, 采样器根据采样网格对待校正图像进行校正, 生成校正后的图像。该过程表示为

$$\bar{I}_m(\mathbf{p}) = I_m(\mathbf{p} + \mathbf{V}(\mathbf{p})) = \sum_{\mathbf{p}' \in N(\mathbf{p} + \mathbf{V}(\mathbf{p}))} I_m(\mathbf{p}') \prod_{d \in \{0,1\}} [1 - |p_d + V_d(\mathbf{p}) - p'_d|]$$

式中: \bar{I}_m 和 I_m 分别是校正后和校正前的待校正图像, \mathbf{p} 和 \mathbf{p}' 均表示像素位置, $\mathbf{V}(\mathbf{p})$ 是位置 \mathbf{p} 处

的运动向量 (由水平分量和垂直分量组成), $N(\mathbf{p} + \mathbf{V}(\mathbf{p}))$ 是位置 $\mathbf{p} + \mathbf{V}(\mathbf{p})$ 的 4-邻域, p_d 和 p'_d 分别是 \mathbf{p} 和 \mathbf{p}' 的第 d 个分量。

1.2 数据集

数据集由临床采集的 IVUS 和 IVOCT 图像序列构成, 用于网络模型的训练和测试。IVUS 图像序列通过配备 2.9 F、30 MHz 机械驱动导管的 Jomed Endosonic (Beringen, Switzerland) 成像系统采集, 导管回撤速度为 0.5 mm/s, B 模式图像的帧采样率是 30 f/s, 图像尺寸为 256×256 像素, 灰度范围为 [0, 255]。IVOCT 图像使用配备 2.7 F C7 Dragonfly OCT 导管的谱域 OCT 系统 (C7XR TM OCT 系统, LightLab Imaging/St Jude Medical Inc, St Paul, MN) 采集, 导管回撤速度为 20 mm/s, 以 120 kHz 的 A 线速率获得 B 模式图像, 图像尺寸为 240 像素×240 像素, 实验中将其调整为 256 像素×256 像素以构建数据集。原始图像序列中的相邻两帧组成一个样本对, 其中前一帧为目标图像, 后一帧为待校正图像。由于相邻帧之间具有较强的相关性, 逐帧校正可以确保图像序列的连续性和一致性, 从而提高整体校正精度。

原始 IVUS 和 IVOCT 数据集中分别包含 8 990 对和 11 990 对样本。将原始样本随机打乱, 按照 6:4 的比例划分为训练集和测试集。IVUS 数据集中, 训练集包含 5 394 对样本, 测试集包含 3 596 对样本。IVOCT 数据集中, 训练集包含 7 194 对样本, 测试集包含 4 796 对样本。为了防止过拟合, 使用随机旋转、随机水平/垂直翻转和随机平移等数据增强技术扩充训练集, 扩充后的 IVUS 和 IVOCT 训练集分别包含 16 182 对和 21 582 对样本图像。测试集中的样本图像未在训练集中出现过。在每个数据集上进行 5 次随机实验, 并采用随机实验的均值评估模型的性能。图 7 是一个 IVOCT 图像序列中样本对的构造示例。

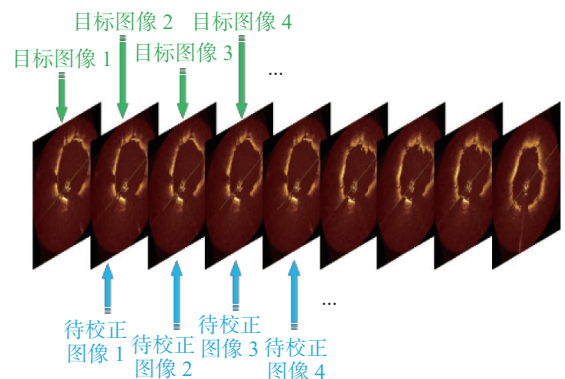


图 7 IVOCT 图像序列中样本对的构造示例

Fig. 7 Sample pair construction in an IVOCT sequence

1.3 优化网络参数

网络的总损失函数定义为

$$\mathcal{L}(\theta) = \eta_1 \mathcal{L}_m(\theta) + \eta_2 \mathcal{L}_d(\theta) + \eta_3 \mathcal{L}_s(\theta)$$

式中： θ 是网络中所有可学习参数的集合， η_1 、 η_2 和 η_3 是超参数，本文实验将其设置为 $\eta_1=1$ 、 $\eta_2=0.01$ 和 $\eta_3=0.05$ 。

\mathcal{L}_m 是光度损失，衡量目标图像和待校正图像之间的亮度差异，以确保校正后的图像与目标图像在亮度上保持一致，定义为

$$\mathcal{L}_m = \sum_p \Psi(I_t(\mathbf{p}) - I_m(\mathbf{p} + \mathbf{V}(\mathbf{p})))$$

式中： $\Psi(\cdot)$ 是鲁棒惩罚函数，定义为 $\Psi(x) = (|x| + \varepsilon)^q$ ，其中 $q=0.4$ ， $\varepsilon=0.01$ ； \mathbf{p} 是像素坐标； \mathbf{V} 是估计出的运动场； $I_t(\mathbf{p})$ 和 $I_m(\mathbf{p})$ 分别是目标图像和待校正图像中像素 \mathbf{p} 处的亮度值。

\mathcal{L}_d 是蒸馏损失，通过逐层比较估计的低分辨率运动场与下采样后的运动场，确保各尺度运动场估计的准确性，定义为

$$\mathcal{L}_d = \sum_{l=1}^5 \sum_p \Psi(\mathbf{V}_l(\mathbf{p}) - \mathcal{S}_l(s_l, \mathbf{V}(\mathbf{p})))$$

式中： \mathbf{V}_l 是第 l 个运动估计模块输出的低分辨率运动场， \mathcal{S}_l 是下采样函数， s_l 是特征提取模块第 l 层的比例放大倍数。

\mathcal{L}_s 是边缘感知平滑损失，用于保证运动场的平滑性，定义为

$$\mathcal{L}_s = \sum_p \Psi(|\partial_x \mathbf{V}(\mathbf{p})| e^{-\alpha|\partial_x I_t(\mathbf{p})|}) + \sum_p \Psi(|\partial_y \mathbf{V}(\mathbf{p})| e^{-\alpha|\partial_y I_t(\mathbf{p})|})$$

式中： α 是边缘权重， ∂_x 和 ∂_y 分别表示 x 方向和 y 方向上的偏导数。

使用 Adam 算法对网络参数进行优化，初始学习率设置为 0.001，等间隔学习率调整的倍数为 0.99，权重衰减为 10^{-4} ，最大 epoch 为 1 000，批量大小为 32。训练过程中，网络损失函数曲线如图 8 所示。

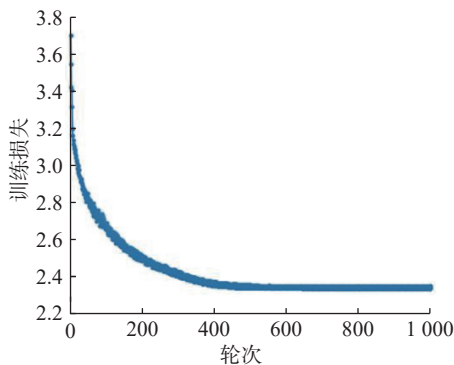


图 8 训练 EIR-Net 的损失函数曲线
Fig. 8 Training loss of EIR-Net

2 方法的性能评价

使用临床采集的 IVUS 和 IVOCT 图像序列验证 EIR-Net 抑制运动伪影的可行性。通过对比抑制伪影前后的纵向视图视觉效果，以及采用平均帧间差异度和标准差作为评价指标，评估对运动伪影的抑制效果。

标准差用于衡量各帧 B 模式图像中从图像中心（即导管中心）到血管壁内膜上边缘的距离的离散程度，定义为

$$s = \sqrt{\frac{1}{N} [(l_1 - \bar{l})^2 + (l_2 - \bar{l})^2 + \dots + (l_N - \bar{l})^2]}$$

式中： N 是图像序列中的总帧数； l_i 是第 i 帧图像中从图像中心到管壁内膜上边缘的距离（单位：像素），如图 9 所示； \bar{l} 是整个图像序列中 l_i ($i=1, 2, \dots, N$) 的平均值。

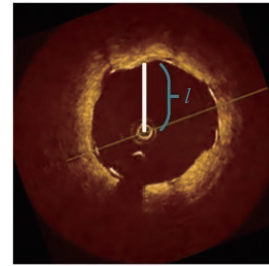


图 9 从图像中心（即导管中心）到血管壁内膜上缘的距离
Fig. 9 Distance from the image center (i.e. the catheter center) to the upper edge of the vessel wall intima

平均帧间差异度是图像序列中相邻帧图像之间特征差异度的平均值：

$$\bar{d} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N d_{i,j}$$

式中 $d_{i,j}$ 是第 i 帧和第 j 帧图像之间的差异值：

$$d_{i,j} = 1 - \text{sim}(\mathbf{v}_i, \mathbf{v}_j)$$

式中 $\text{sim}(\mathbf{v}_i, \mathbf{v}_j)$ 是第 i 帧和第 j 帧图像的特征向量 \mathbf{v}_i 和 \mathbf{v}_j 之间的余弦相似度：

$$\text{sim}(\mathbf{v}_i, \mathbf{v}_j) = \frac{\mathbf{v}_i^T \mathbf{v}_j}{\|\mathbf{v}_i\| \cdot \|\mathbf{v}_j\|} \quad (2)$$

式中 $\|\cdot\|$ 表示向量求模。

以基于深度学习的离线门控方法 IBG-Net^[15]、基于 AP 聚类的门控法^[13]、基于图像配准的非学习方法^[21]以及基于图像配准的深度学习方法 MAC-Net^[32]作为基线方法，与 EIR-Net 进行比较，以验证其优势。IBG-Net 采用无监督学习框架，使用 1.2 节所述的临床图像序列构建训练集和测试集。MAC-Net 使用仿真数据集训练，数据集构建方法参见文献^[32]，测试集由 1.2 节所述的临床

图像序列构成。

本文实验所用的计算机配置为 12th Gen Intel (R) Core (TM) i7-12700H CPU、16 GB RAM 和 Windows 11 64 位操作系统, GPU 为 NVIDIA Tesla P100-PCIE-16 GB, 主频 1.328 5 GHz。网络搭建和训练使用的编程语言为 Python 3.7, 深度学习框架为 PyTorch 1.7。实现基于图像配准的非学习方法和基于 AP 聚类的门控法的编程环境为 Matlab(R2016a, the Math Works, Inc., Natick, Massachusetts)。

3 实验结果与分析

支架植入术是目前治疗冠心病的主要手术方法, 冠状动脉内成像对于指导支架植入、评估手术效果及术后随访具有重要意义。本文采用

临床采集的支架植入前后的 IVUS/OCT 图像数据测试训练后的网络性能, 并对实验结果进行分析。

3.1 支架植入前的临床图像序列的实验结果

采用 EIR-Net 对支架植入之前的 IVUS/OCT 图像序列中的相邻帧进行配准, 部分结果如图 10 所示。为了直观展示运动场, 使用 HSV 颜色编码将运动向量场转换成伪彩色图像, 其中运动向量的方向映射为色调分量, 向量大小映射为饱和度分量, 较大的向量对应更高的饱和度和纯度。由图 10 可知, 经过配准的待校正图像在保留原有结构信息的基础上, 其管腔轮廓与目标图像高度匹配。表 1 列出了运动校正前后的帧间相似度(由式 (2) 计算所得), 可以看出与校正前相比, 目标图像与校正后图像之间的相似度明显提高。

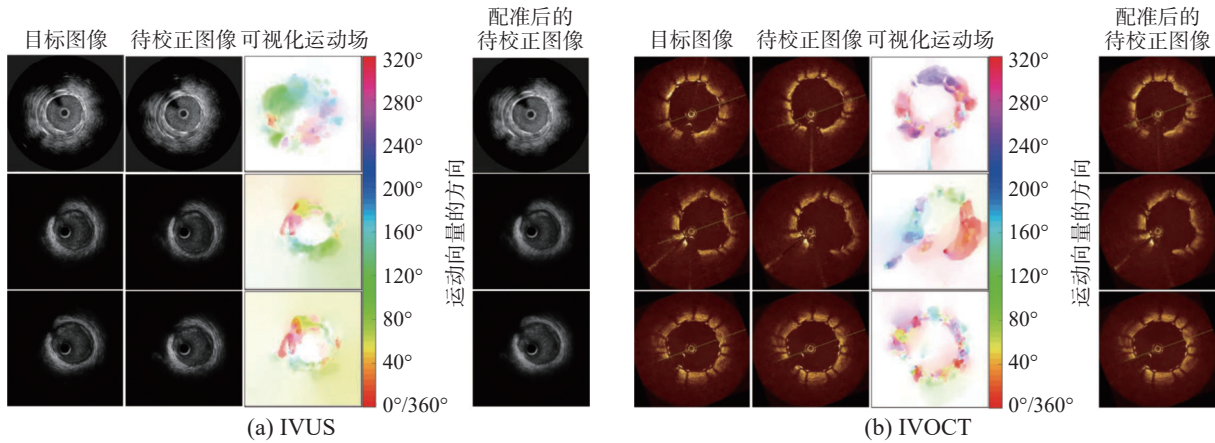


图 10 采用 EIR-Net 对 B 模式 IVUS/OCT 图像进行运动校正的结果

Fig. 10 Results of motion correction for B-model IVUS/OCT images using EIR-Net

表 1 IVUS/OCT 图像序列中相邻两帧图像运动校正前后的帧间相似度

Table 1 Interframe similarity between adjacent images in IVUS/OCT image sequence before and after correction

配准前/后	IVUS			IVOCT		
	样本1	样本2	样本3	样本1	样本2	样本3
配准前	0.939 9	0.875 2	0.917 1	0.907 7	0.872 1	0.944 8
配准后	0.984 5	0.953 9	0.971 6	0.942 5	0.906 6	0.972 8

图 11 是采用 EIR-Net 对 IVUS/OCT 图像序列抑制运动伪影前后的纵向视图和血管三维空间图。可以看出, 抑制伪影后, 图像序列的纵向视图和三维血管图中管壁的平滑度显著提升。图 12 是采用不同方法抑制运动伪影前后 IVUS/OCT 图像序列的平均帧间差异度和标准差对比。与抑制之前相比, 抑制伪影后图像序列的平均帧间差异度和标准差均显著下降。同时, EIR-Net 抑制运动

伪影后的图像序列的平均帧间差异度和标准差明显优于基线方法。这是由于 EIR-Net 通过上采样模块和运动估计模块的协同工作, 实现对运动场的逐步细化。而且, 在运动估计模块中引入了代价体, 以捕捉相邻帧之间的特征匹配信息, 从而提供高精度的运动场, 将运动场作用于待校正图像上, 使得校正后图像与目标图像相似度更高, 运动伪影抑制效果更好。

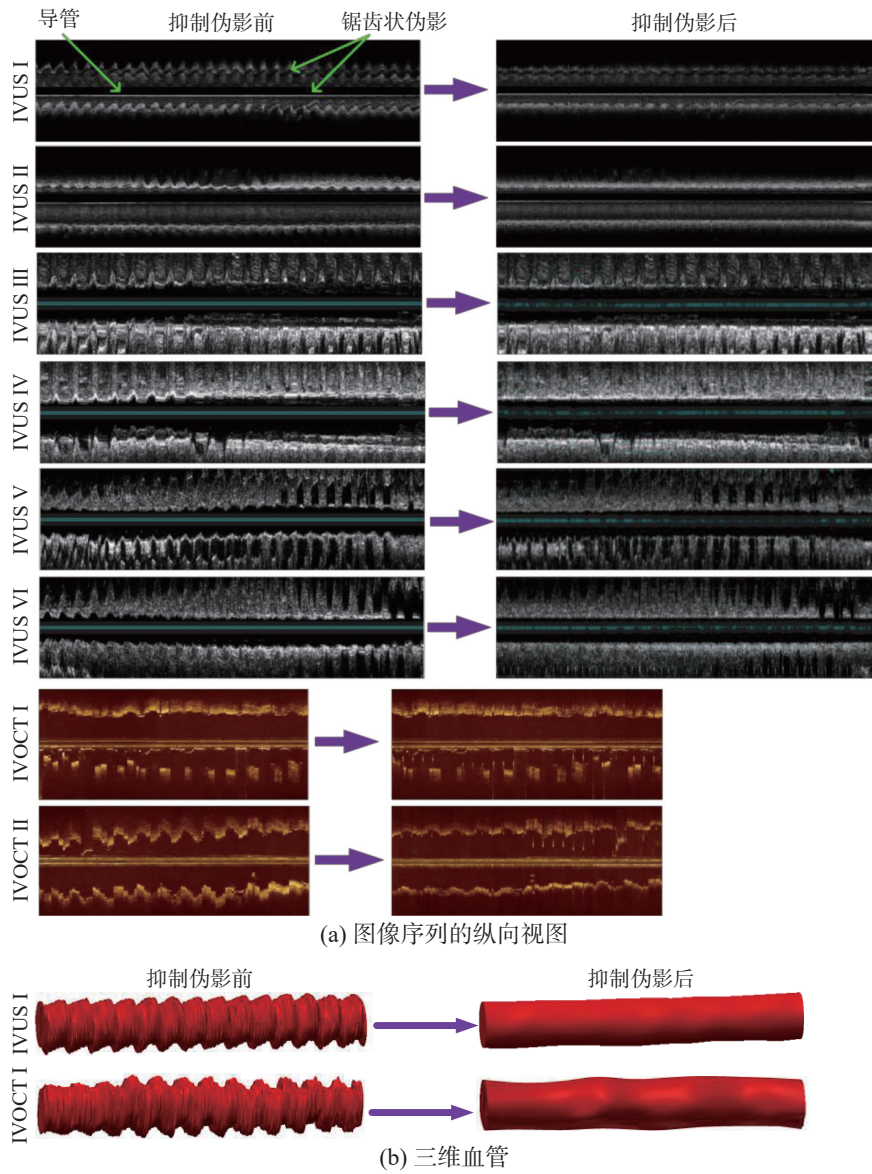
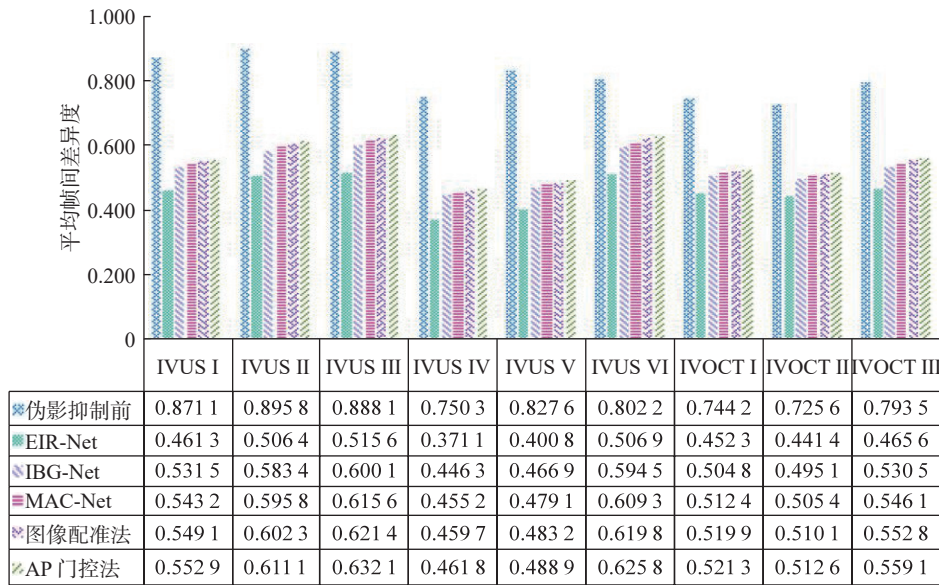


图 11 采用 EIR-Net 对 IVUS/OCT 图像序列进行运动伪影抑制的结果

Fig. 11 Results of motion artifact suppression using EIR-Net on IVUS/OCT image sequences



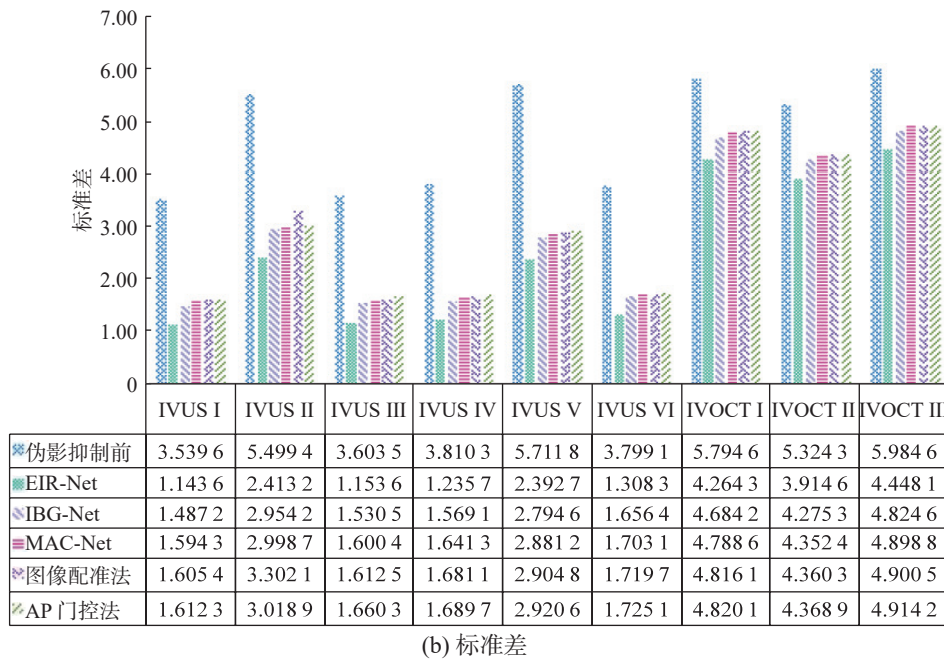


图 12 采用不同方法抑制运动伪影前后 IVUS/OCT 图像序列的平均帧间差异度和标准差对比

Fig. 12 Comparison of average interframe dissimilarity and standard deviation of IVUS/OCT image sequences before and after motion artifact suppression

表 2 记录了不同方法的处理时间, 其中 3 种深度学习方法的处理时间不包括训练时长。MAC-Net 需要利用 AP 聚类对图像序列进行静态和动态帧的划分, 而基于图像配准的非学习方法需要对各帧图像进行分割以提取血管壁边缘, 因此两者的处理时间较长, 很难满足实时性要求。EIR-Net、AP 门控法和 IBG-Net 均采用端到端方

式处理图像序列, 因此处理时间较短。相比之下, IBG-Net 的网络模型更简单, 计算效率略高, 但 IBG-Net 和 AP 门控法需要丢弃大量帧, 可能造成血管信息的丢失。EIR-Net 在保证图像序列完整性的同时抑制了运动伪影, 适用于血管三维重建、血流动力学评估和生物力学特性分析等应用场景。

表 2 不同方法的 GPU 运行时间
Table 2 GPU runtime of different methods

图像序列	序列长度/f	GPU运行时间/s				
		EIR-Net	IBG-Net	MAC-Net	图像配准法	AP门控法
IVUS I	600	54.24	15.31	675.83	727.01	91.30
IVUS II	600	53.96	15.19	676.56	726.37	90.60
IVUS III	600	53.11	15.26	677.03	726.85	91.70
IVOCT	1 200	90.83	8.29	1 074.20	1 171.16	51.10

在对临床 IVUS/OCT 图像序列进行运动伪影抑制的研究中, 由于无法获知血管段体积的真实值, 因此通常采用 ECG 门控的结果作为金标准^[34]。表 3 和图 13 给出了采用不同方法抑制伪影前后

图像序列中血管段的体积对比情况。可以看出, 采用 EIR-Net 得到的血管体积与 ECG 门控法的体积最接近, 表明 EIR-Net 能够在抑制伪影的同时准确量化血管段体积并达到了较高的精度。

表 3 抑制运动伪影前后 IVUS/OCT 图像序列中血管段的体积
Table 3 Volume of vascular segments in IVUS/OCT image sequences before and after suppressing motion artifacts

图像序列	序列长度/f	血管段体积/mm ³						
		抑制伪影前	ECG门控	EIR-Net	IBG-Net	MAC-Net	图像配准法	AP门控法
IVUS I	600	265.453	246.962	248.143	251.842	243.631	240.916	239.764

续表 3

图像序列	序列长度/f	血管段体积/mm ³						
		抑制伪影前	ECG门控	EIR-Net	IBG-Net	MAC-Net	图像配准法	AP门控法
IVUS II	600	269.369	258.365	257.137	254.567	263.084	265.272	265.412
IVUS III	600	295.024	289.227	285.593	284.759	283.624	282.239	282.116
IVUS IV	300	116.003	101.736	97.891	96.473	94.340	93.087	92.848
IVUS V	300	142.672	131.880	126.841	125.664	124.158	122.937	122.454
IVUS VI	300	139.505	124.436	128.613	128.564	129.701	119.011	118.775
IVOCT I	1 200	529.816	512.068	516.208	517.112	518.271	519.372	519.654
IVOCT II	1 200	551.003	536.794	534.579	532.722	541.903	542.615	543.615
IVOCT III	1 200	551.683	528.982	531.592	533.226	536.538	537.724	538.046

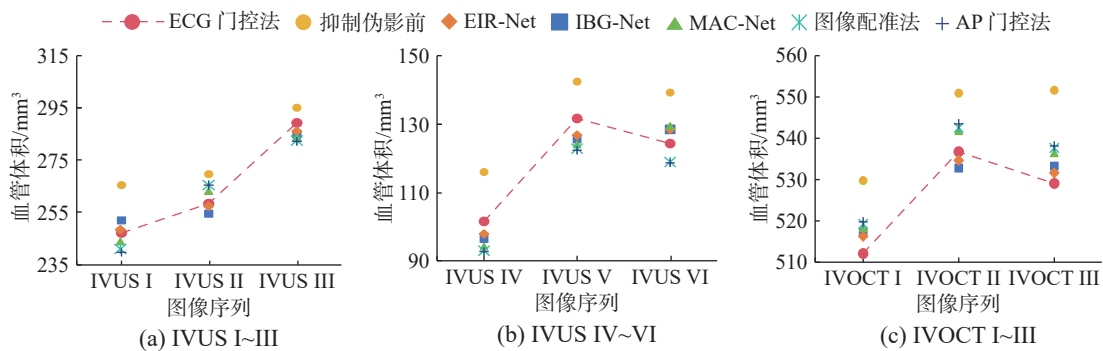


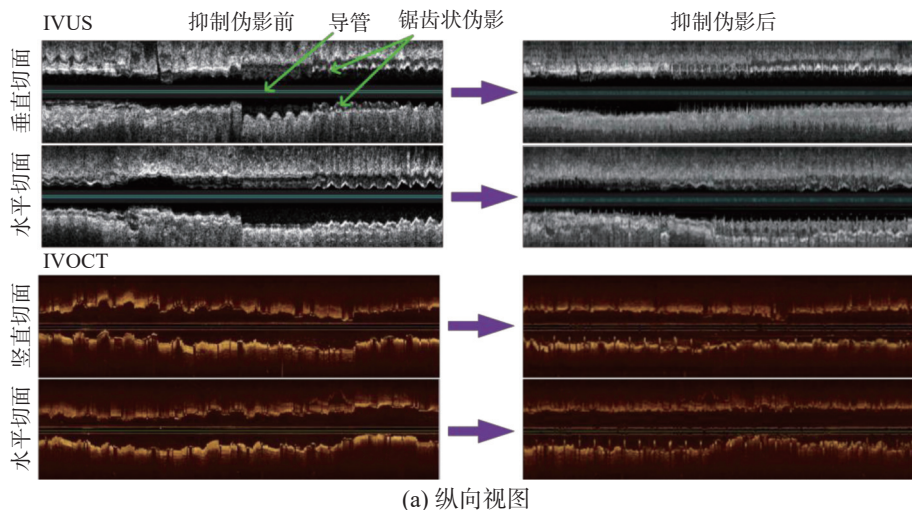
图 13 采用不同方法抑制运动伪影之后 IVUS/OCT 图像序列中血管段体积的比较

Fig. 13 Comparison of vascular segment volumes in IVUS/OCT image sequences after suppressing motion artifacts using different methods

3.2 支架植入术后图像序列的实验结果

常用的心脏支架是可弯曲的金属网格，支架的弹性与血管壁的弹性存在差异，心脏的周期性运动对支架的影响相对较小。因此，传统的门控法在抑制支架植入术后图像序列中的运动伪影

影方面效果有限^[15]。采用 EIR-Net 对支架植入术后图像序列的运动伪影抑制结果如图 14 所示。可以看出，抑制伪影后，纵向视图中的血管壁边缘变得平滑，平均帧间差异度和标准差也显著降低。



(a) 纵向视图

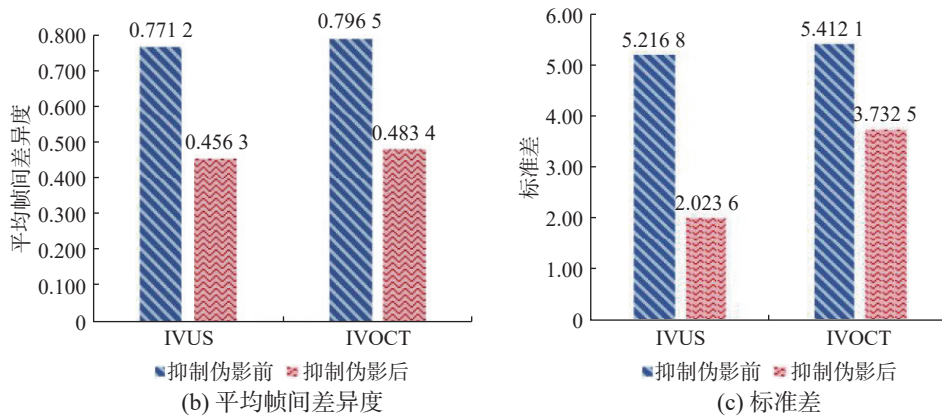


图 14 采用 EIR-Net 对支架植入术后的 IVUS/OCT 图像序列抑制运动伪影的结果

Fig. 14 Results of using EIR-Net to suppress motion artifacts in IVUS/OCT image sequences after stent implantation

3.3 消融实验结果

本节结合对临床图像的实验结果, 分析 EIR-Net 中蒸馏损失和上采样模块两个重要组件对方法性能的影响, 并讨论所提方法相对于其他运动伪影抑制方案的优势和不足。

设计如下消融实验: 分别去除 EIR-Net 中的上采样模块和蒸馏损失, 得到的网络模型分别命名为 Loss-model 和 Sgu-model。从训练损失和图像配准结果两个方面对比分析 Loss-model 和 Sgu-model 与 EIR-Net 的表现。图 15 是各消融实验模型的训练损失曲线。可以看出, Sgu-model 的收敛速度较慢且收敛效果不佳。Loss-model 的训练损失震荡较大, 这是由于 Loss-model 不包含上采样模块, 尽管降低了网络模型的复杂性, 但限制了模型学习信息的能力, 导致训练过程的不稳定。相比之下, EIR-Net 的收敛速度快且效果明显优于两个消融实验模型。

图 16 是分别采用 EIR-Net、Loss-model 和 Sgu-model 对两帧 IVUS/OCT 图像进行运动校正

的结果, 图中用方框标注了部分细节。可以看出, EIR-Net 校正后的图像与目标图像的管腔轮廓更加一致, 细节更为清晰。Loss-model 的校正后图像与目标图像的管腔轮廓基本一致, 但细节较模糊。Sgu-model 的校正后图像与目标图像中的管腔轮廓在细节和角度上均有差异。

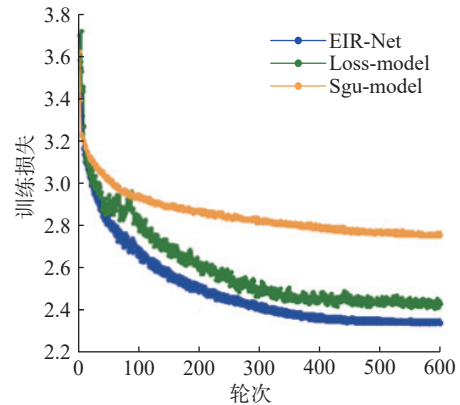
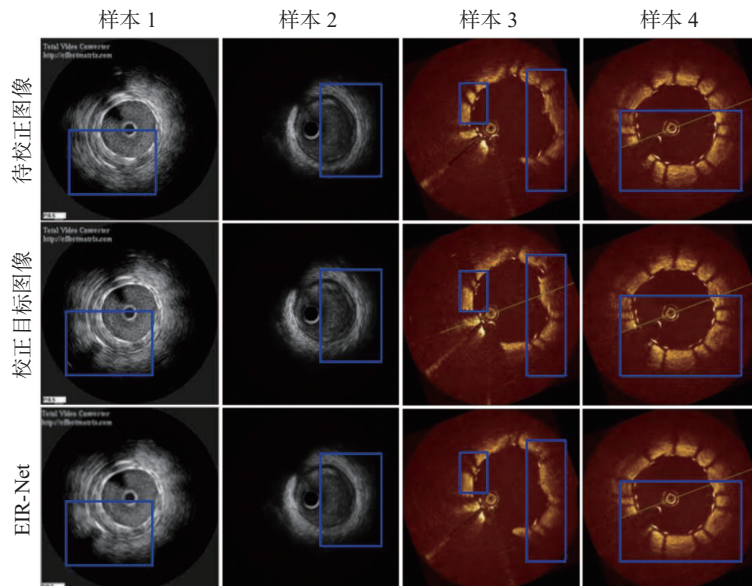


图 15 EIR-Net、Loss-model 和 Sgu-model 的训练损失曲线
Fig. 15 Training losses of EIR-Net, Loss-model and Sgu-model



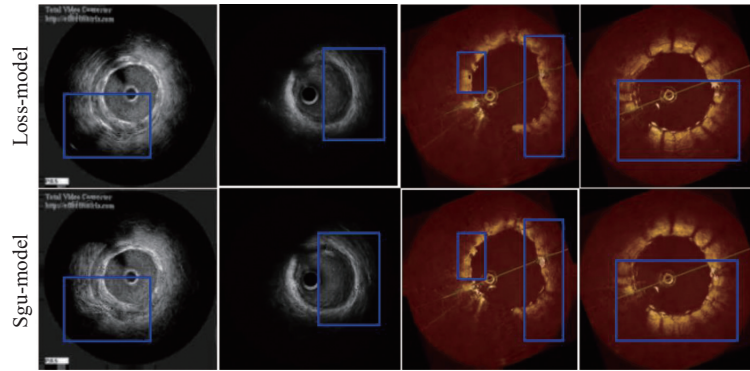


图 16 分别采用 EIR-Net、Loss-model 和 Sgu-model 进行图像配准的结果

Fig. 16 Results of image registration using EIR-Net, Loss-model, and Sgu-model, respectively

连续回撤导管采集图像的数据量非常大,在回撤速度为 1 mm/s 且帧采样率为 20 f/s 的情况下,对约 50 mm 长的血管段一次回撤导管即可获得约 1000 f 图像。因此,相邻两帧图像之间的内容差别很小,难以从视觉上区分血管轮廓的位置、大小和角度等细节变化。本文采用余弦相似度作为评价指标,对各消融实验模型的配准结果进行定量评价,结果如表 4 所示。显然与 Loss-model 和 Sgu-model 相比, EIR-Net 校正运动后的图像与目标图像的特征向量之间的余弦相似度更高,证明当上采样模块和蒸馏损失共同作用时,可以显著提高精度。

表 4 分别采用 EIR-Net、Loss-model 和 Sgu-model 进行图像配准前后的余弦相似度

Table 4 Cosine similarity before and after image registration using EIR-Net, Loss-model, and Sgu-model, respectively

样本名称	配准前	EIR-Net	Loss-model	Sgu-model
样本1	0.939 9	0.984 5	0.953 3	0.906 1
样本2	0.875 2	0.953 9	0.902 1	0.863 8
样本3	0.872 1	0.906 6	0.881 4	0.859 9
样本4	0.944 8	0.972 8	0.958 9	0.929 6

4 讨论

针对 IVUS/OCT 图像序列中运动伪影的直接抑制和补偿问题,本文提出了一种无需依赖预先标记训练数据的解决方案。采用临床图像序列对网络模型进行无监督训练,有效地识别运动场,克服了现有模型对运动场已知的仿真数据集的高度依赖性。该方法不仅提高了模型处理大规模未标记 IVUS/OCT 图像数据集的能力,而且降低了对标记数据的依赖。

采用光流法进行运动估计和校正时,需要基于亮度恒定、空间一致和时间连续 3 个假设。然

而在介入血管内成像过程中,由于导管的快速移动、血管的不规则形态和复杂的血流模式,相邻帧图像记录的并不是同一血管横截面,因此这些假设很难同时满足。EIR-Net 以图像序列的首帧作为目标图像,后续每帧图像以校正后的前一帧作为目标图像进行配准,这种方式存在误差传递和累积的问题。帧数越多,误差也越大,表现为横截面图像的模糊和失真,以及纵向视图中细节的模糊和丢失,使得 EIR-Net 不适合用于较长图像序列中的运动伪影抑制。未来的工作可以考虑在运动估计之前对图像进行去噪预处理,并对每次配准后的待校正图像进行增强,以减少伪影抑制过程中的误差积累。

在支架植入术后的图像中,支架支脚会遮挡部分血管组织,被遮挡区域中的像素会导致运动场的计算误差,进而影响图像配准精度。未来的工作可以考虑采用自动图像分割技术标记支架遮挡区域为无效区域,生成相应掩码,并将掩码应用于损失函数中,使网络模型在训练过程中不受支架遮挡区域的影响,以提高运动估计精度。

此外,未来将探索集成自监督学习等先进深度学习技术,增强无监督环境下的模型泛化和特征提取;同时计划实施多模态学习,整合 IVUS 和 OCT 数据,利用互补信息优化运动伪影抑制;还将优化网络结构,引入注意力机制和深度可分离卷积,降低计算负担,以适应实时临床需求。

5 结束语

本文提出了一种抑制冠状动脉内图像序列中运动伪影的深度学习方法。所搭建的深度神经网络 EIR-Net 通过估计相邻两帧横截面图像之间的运动场,并对各帧图像进行运动校正,达到抑制运动伪影的目的。利用临床采集的 IVUS 和 IVOCT 图像序列对网络进行的无监督训练和测试表明, EIR-Net 能在保持图像序列数据完整性的

同时,有效抑制运动伪影。与原始图像序列相比,处理后的纵向视图视觉效果显著提升,而且平均帧间差异度和标准差等定量评价指标均表现出明显改进。对比实验表明,EIR-Net抑制运动伪影的效果整体优于IBG-Net、MAC-Net、AP门控法以及基于图像配准的非学习方法。消融实验进一步表明,当EIR-Net中蒸馏损失和上采样模块共同作用时,运动校正精度显著提升。

参考文献:

- [1] SUWANNASOM P, CHICHAREON P, ROONGSANG-MANOON W, et al. Impact of the adjunctive use criteria for intravascular ultrasound-guided percutaneous coronary intervention and clinical outcomes[J]. *Scientific reports*, 2023, 13(1): 711–711.
- [2] MOHAN N C, JOHNSON T W. Intracoronary optical coherence tomography—an introduction[J]. *Catheterization and cardiovascular interventions*, 100(S1): S57–S65.
- [3] LENG J, ZHANG J, LI C, et al. Multi-spectral intravascular photoacoustic/ultrasound/optical coherence tomography tri-modality system with a fully-integrated 0.9-mm full field-of-view catheter for plaque vulnerability imaging[J]. *Biomedical optics express*, 2021, 12(4): 1934–1946.
- [4] KERMANI A, AYATOLLAHI A, TAKI A. Novel coronary tortuosity extraction using intravascular ultrasound sequence alone[C]//IEEE 2017 24th National and 2nd International Iranian Conference on Biomedical Engineering. Tehran: IEEE, 2017: 47–50.
- [5] NIKOPOULOS S, PAPAFAKLIS MI, TSOMPOU P, et al. Virtual hemodynamic assessment of coronary lesions: the advent of functional angiography and coronary imaging[J]. *Journal of clinical medicine*, 2024, 13(8): 2243.
- [6] TIAN Jinwei, PENG Xiang, ZHANG Yanwen, et al. In vivo detection of cholesterol crystals in atherosclerotic plaque with optical coherence tomography[C]//Cholesterol Crystals in Atherosclerosis and Other Related Diseases, Contemporary Cardiology. New York: Springer, 2023: 75–87.
- [7] KYRIAKIDIS S, RIGAS G, KIGKA V, et al. An all-in-one tool for 2D atherosclerotic disease assessment and 3D coronary artery reconstruction[J]. *Journal of cardiovascular development and disease*, 2023, 10(3): 130.
- [8] GAUR P, GRIECO A, FAINMAN Y. Multirate spectral domain optical coherence tomography[J]. *IEEE photonics journal*, 2023, 15(5): 1–10.
- [9] JENSEN L O, THAYSSSEN P. Accuracy of electrocardiographic-gated versus nongated volumetric intravascular ultrasound measurements of coronary arterial narrowing[J]. *The American journal of cardiology*, 2007, 99(2): 279–283.
- [10] CECCHETTI L, WANG T, HOOGENDOORN A, et al. In-vitro and in-vivo imaging of coronary artery stents with Heartbeat OCT[J]. *International journal of cardiovascular imaging*, 2020, 36: 1021–1029.
- [11] XIA Jun, CHEN Wanyi, MASLOV K, et al. Retrospective respiration-gated whole-body photoacoustic computed tomography of mice[J]. *Journal of biomedical optics*, 2014, 19(1): 016003.
- [12] 孙正. 抑制冠状动脉内超声图像序列运动伪影的离线门控方法[J]. *光电子·激光*, 2010, 21(4): 632–638.
- [13] SUN Zheng, DU Jiejie. An off-line gating method for suppressing motion artifacts in intracoronary ultrasound image sequence[J]. *Journal of optoelectronics-laser*, 2010, 21(4): 632–638.
- [14] SUN Zheng, DU Jiejie. Suppression of motion artifacts in intravascular photoacoustic image sequences[J]. *Biomedical optics express*, 2021, 12(11): 6909–6927.
- [15] BAJAJ R, HUANG Xingru, KILIC Y, et al. A deep learning methodology for the automated detection of end-diastolic frames in intravascular ultrasound images[J]. *International journal of cardiovascular imaging*, 2021, 37(6): 1825–1837.
- [16] SUN Zheng, YAO Yue, WANG Ru. An unsupervised deep learning framework for retrospective gating of catheter-based cardiac imaging[J]. *IET signal processing*, 2024: 5664618.
- [17] DANILOUCHKINE M G, MASTIK F, VAN DER STEEN A F. A study of coronary artery rotational motion with dense scale-space optical flow in intravascular ultrasound[J]. *Physics in medicine and biology*, 2009, 54(6): 1397–1418.
- [18] DANILOUCHKINE M G, MASTIK F, VAN DER STEEN A F. Improving IVUS palpography by incorporation of motion compensation based on block matching and optical flow[J]. *IEEE transactions on ultrasonics ferroelectrics and frequency control*, 2008, 55(11): 2392–2404.
- [19] SUN Deqing, YANG Xiaodong, LIU Mingyu, et al. PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8934–8943.
- [20] LUO Kunming, WANG Chuan, LIU Shuaicheng, et al. UPFlow: upsampling pyramid for unsupervised optical flow learning[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 1045–1054.
- [21] 张聪炫, 史世栋, 葛利跃, 等. 基于遮挡优化的金字塔块匹配光流估计方法[J]. *电子学报*, 2023, 51(9): 2539–2548.
- [22] ZHANG Congxuan, SHI Shidong, GE Liyue, et al. Pyr-

- amid patch-matching optical flow estimation method based on occlusion optimization[J]. *Acta electronica sinica*, 2023, 51(9): 2539–2548.
- [21] TALOU G D M, BLANCO P J, LARRABIDE I, et al. Registration methods for IVUS: transversal and longitudinal transducer motion compensation[J]. *IEEE transactions on biomedical engineering*, 2017, 64(4): 890–903.
- [22] SUN Zheng, WANG Jianjian. Compensation of in-plane rigid motion for in vivo intracoronary ultrasound image sequence[J]. *Computers in biology and medicine*, 2013, 43(9): 1077–1085.
- [23] 余升林, 吴彤, 葛明锋, 等. 循环一致性的无监督可变形图像配准方法[J]. *计算机辅助设计与图形学学报*, 2023, 35(4): 516–524.
YU Shenglin, WU Tong, GE Mingfeng, et al. Unsupervised deformable image registration method with cyclic consistency[J]. *Journal of computer-aided design & computer graphics*, 2023, 35(4): 516–524.
- [24] FU Yunguan, BROWN N M, SAEED S U, et al. DeepReg: a deep learning toolkit for medical image registration[J]. *Journal of open source software*, 2020, 5(55): 2705.
- [25] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial Transformer networks[C]//*Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: MIT Press, 2015: 2017–2025.
- [26] GÜLER A R, TRIGEORGIS G, ANTONAKOS E, et al. DenseReg: fully convolutional dense shape regression in-the-wild[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 6799–6808.
- [27] 赵华秋, 谢勤岚. 一种用于心脏运动估计的快速图像配准方法[J]. *中南民族大学学报(自然科学版)*, 2023, 42(1): 120–127.
ZHAO Huaqiu, XIE Qinlan. A fast image registration method for cardiac motion estimation[J]. *Journal of South-central Minzu University(Natural Science Edition)*, 2023, 42(1): 120–127.
- [28] 黄鹏, 郭春生, 陈华华, 等. 基于深度学习的图像配准方法综述[J]. *杭州电子科技大学学报(自然科学版)*, 2020, 40(6): 37–44.
HUANG Peng, GUO Chunsheng, CHEN Huahua, et al. An overview of image registration methods based on depth learning[J]. *Journal of Hangzhou Dianzi University(Natural Sciences)*, 2020, 40(6): 37–44.
- [29] 陈建明, 曾祥津, 钟丽云, 等. 基于深度学习的图像配准方法研究进展[J]. *量子电子学报*, 2022, 39(6): 899–926.
CHEN Jianming, ZENG Xiangjin, ZHONG Liyun, et al. Research progress of image registration methods based on deep learning[J]. *Chinese journal of quantum electronics*, 2022, 39(6): 899–926.
- [30] 应时辉, 杨苑, 杜少毅, 等. 基于深度学习的医学影像配准综述[J]. *模式识别与人工智能*, 2021, 34(4): 287–299.
YING Shihui, YANG Wan, DU Shaoyi, et al. Deep learning based medical image registration: a review[J]. *Pattern recognition and artificial intelligence*, 2021, 34(4): 287–299.
- [31] 周勤, 王远军. 基于深度学习的医学图像配准研究进展[J]. *上海理工大学学报*, 2021, 43(5): 421–428.
ZHOU Qin, WANG Yuanjun. Advances in medical image registration based on deep learning[J]. *Journal of University of Shanghai for Science and Technology*, 2021, 43(5): 421–428.
- [32] SUN Zheng, DU Jiejie, YAO Yue, et al. A deep learning method for motion artifact correction in intravascular photoacoustic image sequence[J]. *IEEE transactions on medical imaging*, 2023, 42(1): 68–78.
- [33] HUR J, ROTH S. Iterative residual refinement for joint optical flow and occlusion estimation[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 5747–5756.
- [34] DE WINTER S A, HAMERS R, DEGERTEKIN M, et al. Retrospective image-based gating of intracoronary ultrasound images for improved quantitative analysis: the intelligate method[J]. *Catheterization and cardiovascular interventions*, 2004, 61(1): 84–94.

作者简介:



王茹, 硕士研究生, 主要研究方向为深度学习和血管内超声/OCT 图像处理技术。E-mail: 1820503691@qq.com。



孙正, 教授, 主要研究方向为医学影像技术、多模态成像技术、图像重建和反问题求解。主持国家自然科学基金项目、中国博士后科学基金项目等 10 余项, 获发明专利授权 30 余项。发表学术论文 100 余篇, 出版学术专著 2 部。E-mail: sunzheng@ncepu.edu.cn。



姚越, 硕士研究生, 主要研究方向为深度学习和心脏图像处理。E-mail: yaoyue1098599943@163.com。

[责任编辑: 丁钰]