

DOI: 10.13382/j.jemi.B2609320

面向鲁棒远程光电容积描记的双分支 时空解耦与重构方法*

张建海^{1,2} 姜 丰^{1,2} 程世超^{1,2} 赵昶辰^{1,2}

(1. 杭州电子科技大学计算机学院 杭州 310018; 2. 全省(浙江)脑机协同智能技术及应用重点实验室 杭州 310018)

摘 要: 远程光电容积描记技术可利用普通摄像设备从人脸视频中恢复脉搏相关生理信号,但面部颜色变化幅度微弱,易受头部运动、姿态变化、光照波动、阴影遮挡和成像噪声等非生理因素干扰。现有深度学习方法多采用端到端回归范式,通常缺乏对脉搏成分与噪声成分生成机制的显式刻画,导致模型中间状态可解释性不足。针对上述问题,提出一种面向鲁棒 rPPG 的物理先验引导双分支时空解耦与重构网络 PNRNet。该方法将面部 RGB 时序变化建模为生理脉搏分量与非生理噪声分量的叠加,分别通过独立 3D 卷积分支学习脉搏相关动态和干扰相关动态,并引入特征正交约束降低两类隐表征的冗余耦合。受二色反射模型启发,进一步设计 RGB 时序重构模块,将预测脉搏信号和噪声信号投影回 RGB 观测域,使网络同时受到参考 PPG 波形监督和视频观测域重构约束。实验结果表明,PNRNet 在 iBVP、UBFC-rPPG 和 PURE 数据集上的平均绝对值(MAE)分别为 3.02、1.18 和 1.78 bpm,其中在 iBVP 数据集上较最佳对比方法降低 8.8%。波形分析进一步显示,脉搏分支输出与真实 PPG 在主周期和整体趋势上保持较好一致性,噪声分支输出能够反映部分运动和光照引起的非生理波动。结果说明,结构化引入物理先验有助于提升 rPPG 模型的鲁棒性和可解释性。

关键词: 远程光电容积描记;物理先验神经建模;脉搏—噪声解耦;RGB 信号重构;时空特征解耦

中图分类号: TN911.73;TP391 **文献标识码:** A **国家标准学科分类代码:** 310.61

Dual-branch spatiotemporal disentanglement and reconstruction for robust remote photoplethysmography

Zhang Jianhai^{1,2} Jiang Feng^{1,2} Cheng Shichao^{1,2} Zhao Changchen^{1,2}

(1. School of Computer Science, Hangzhou Dianzi University, Hangzhou 310018, China; 2. Zhejiang Provincial Key Laboratory of Brain Computer Collaborative Intelligence Technology and Applications, Hangzhou 310018, China)

Abstract: Remote photoplethysmography (rPPG) can recover pulse-related physiological signals from facial videos captured by ordinary cameras. However, the cardiac-induced color variations are extremely subtle and are easily corrupted by head motion, pose variation, illumination fluctuation, shadow occlusion, and imaging noise. Most existing deep learning methods formulate rPPG estimation as an end-to-end regression problem, but they usually lack explicit modeling of the generation mechanisms of pulse and noise components, making the intermediate representations difficult to interpret. To address this problem, this paper proposes PNRNet, a physics-guided dual-branch spatiotemporal disentanglement and reconstruction network for robust rPPG. The proposed method models facial RGB temporal variations as a superposition of physiological pulse components and non-physiological noise components. Two independent 3D convolutional branches are used to learn pulse-related and interference-related dynamics, and a feature orthogonality constraint is introduced to reduce redundant coupling between the two latent representations. Inspired by the dichromatic reflection model, an RGB temporal reconstruction module projects the predicted pulse and noise signals back to the RGB observation domain, enabling the network

收稿日期:2026-04-30 Received Date: 2026-04-30

* 基金项目:国家自然科学基金面上项目(62573171)、浙江省自然科学基金项目(LZ25F030005,LY24F020015)、浙江省属高校基本科研业务费专项资金(GK269910299001-009)、全省(浙江)脑机协同智能技术及应用重点实验室(含中央引导地方科技发展资金:2025ZY01045、2026ZY01008;计划编号:2025E10015)项目资助

to be jointly supervised by reference PPG waveforms and video-domain reconstruction. Experimental results show that PNRNet achieves MAE of 3.02, 1.18, and 1.78 bpm on the iBVP, UBFC-rPPG, and PURE datasets, respectively, with an 8.8% MAE reduction over the best competing method on iBVP. These results indicate that the proposed pulse-noise disentanglement and RGB observation-domain reconstruction mechanism helps improve the robustness and interpretability of rPPG models under complex interference conditions.

Keywords: remote photoplethysmography; physics-guided neural modeling; pulse-noise disentanglement; RGB signal reconstruction; spatiotemporal feature disentanglement

0 引言

对生理参数进行连续、准确监测是智慧医疗、远程医疗、运动健康管理和人机交互等场景中的基础需求。传统生理参数测量多依赖心电电极、指夹式脉搏血氧仪等接触式或佩戴式设备,虽然具有较好的测量稳定性,但通常需要与人体皮肤直接接触,在长期佩戴、睡眠监测、新生儿监护和公共健康筛查等场景中可能带来不适、感染风险或操作成本。远程光电容积描记技术(remote photoplethysmography, rPPG)能够利用普通摄像设备从人脸视频中恢复脉搏相关生理信号,具有非接触、低成本和便于部署等特点,因此逐渐成为视觉生理测量领域的研究方向之一。相关研究表明,基于视觉的非接触式生理感知已由心率监测逐步扩展到呼吸率、心率变异性、血压估计等多类生理参数测量^[1-4]。

rPPG 的基本原理是心脏搏动会引起皮下血容量周期性变化,进而改变皮肤对入射光的吸收和反射特性,使普通 RGB 视频中包含与脉搏相关的微弱颜色变化。早期研究验证了从普通 RGB 视频绿色通道恢复脉搏波的可行性^[5],光学远程心搏检测研究也说明,非接触视频信号中存在可由心搏调制的光学变化^[6]。从成像机制看,二色反射模型指出,皮肤 RGB 观测由扩散反射、镜面反射、光源强度调制和成像噪声共同决定^[7]。其中,血容量变化主要影响扩散反射中的脉动成分,而头部运动、表情变化、姿态改变、局部阴影、环境光波动、相机自动曝光和压缩噪声等因素会改变面部区域的 RGB 时序变化,使脉搏相关成分与非生理扰动在观测空间中发生混合。因此,如何从混合 RGB 时序中稳定分离脉搏相关信息,是 rPPG 信号提取需要解决的核心问题^[8]。

传统 rPPG 方法主要基于皮肤光学模型、颜色空间投影和手工信号处理流程。Poh 等^[9]利用盲源分离从面部区域颜色变化中估计心率;CHROM 通过构造色度信号削弱镜面反射和运动干扰^[10];PBV 利用血容量脉搏在 RGB 通道中的特征方向提高运动鲁棒性^[11];SSR 从空间子空间旋转角度建模脉搏相关变化^[12];粒子群算法(particle swarm optimization, PSO)进一步总结了远程 PPG 中颜色投影、归一化和滤波的基本假设^[13]。国内研究也从普通摄像头心率测量^[14]、非合作面部晃动条件下

的心率检测^[15]、基于主成分分析和重叠直方图统计的视频心率测量^[16],以及人脸检测、跟踪和 ROI 提取^[17]等方面进行了探索。这类方法通常具有较明确的物理含义和较低的计算代价,但其性能较依赖固定颜色投影或特定统计假设。当面部运动、姿态变化和光照扰动较为复杂时,假设可能难以充分适应不同采集条件下的 RGB 时序变化。

随着深度学习的发展,研究者开始直接从视频片段中学习 rPPG 相关的时空特征。DeepPhys、PhysNet 和 BVPNet 等方法分别从注意力建模、时空卷积和视频到 BVP 映射等角度提升了非接触式脉搏恢复能力^[18-20];多视角神经网络、JAMSNet、GLISNet 和 WTC3D 等进一步利用多源表征、联合注意力、多尺度融合或高效时空网络增强复杂场景下的建模能力^[21-24]。与传统颜色投影方法相比,深度模型能够自适应学习局部颜色变化和时间动态特征,但多数方法仍主要关注从输入视频到目标脉搏波形的直接映射,对脉搏成分与运动、光照、阴影、传感器噪声等非生理成分之间关系的显式刻画仍不充分,容易出现 rPPG 输出与噪声扰动耦合的问题。

基于上述分析,本文提出一种物理先验引导的双分支时空解耦与 RGB 时序重构框架 PNRNet (pulsenoise-reconstruct-net)。PNRNet 使用两个参数独立的 3D 卷积分支分别学习生理脉搏动态和非生理扰动动态,并通过特征正交约束降低两类隐表征的冗余耦合。进一步地,本文将预测脉搏信号与噪声信号投影回 RGB 观测域,并以 ROI 均值去除和归一化后的 RGB 时间序列作为重构监督,使网络不仅预测参考 PPG 波形,还需要解释视频观测中的主要时序变化。该设计旨在提升模型在复杂干扰条件下的鲁棒性,并增强中间表征的可解释性。

本文提出面向 rPPG 的物理先验双分支解耦框架,将面部时序颜色变化显式拆分为生理脉搏分量与非生理噪声分量,设计了 RGB 观测域重构机制,把二色反射模型启发的信号形成先验嵌入网络训练,使脉搏分支与噪声分支共同解释原始 RGB 时序变化,保留并输出中间噪声表征,使研究者能够审查非生理扰动是否符合运动、光照、阴影和成像噪声等物理先验,在 iBVP、UBFC-rPPG 和 PURE 数据集上验证所提方法在误差、信噪比和波形稳定性方面的有效性。

1 方法

1.1 总体框架

给定输入面部视频片段 $X \in \mathbb{R}^{B \times 3 \times T \times H \times W}$, 其中 B 为批大小, T 为时间长度, H 和 W 为空间分辨率。PNRNet 的目标不是直接从混合观测中回归单一脉搏波形, 而是将面部 RGB 时序变化分解为生理脉搏分量和非生理噪声分量。整体框架由脉搏分支、噪声分支、特征正交约束、RGB 时序重构模块和联合损失函数组成。

脉搏分支输出预测脉搏信号 $\hat{s}_p$ 和脉搏相关隐特征

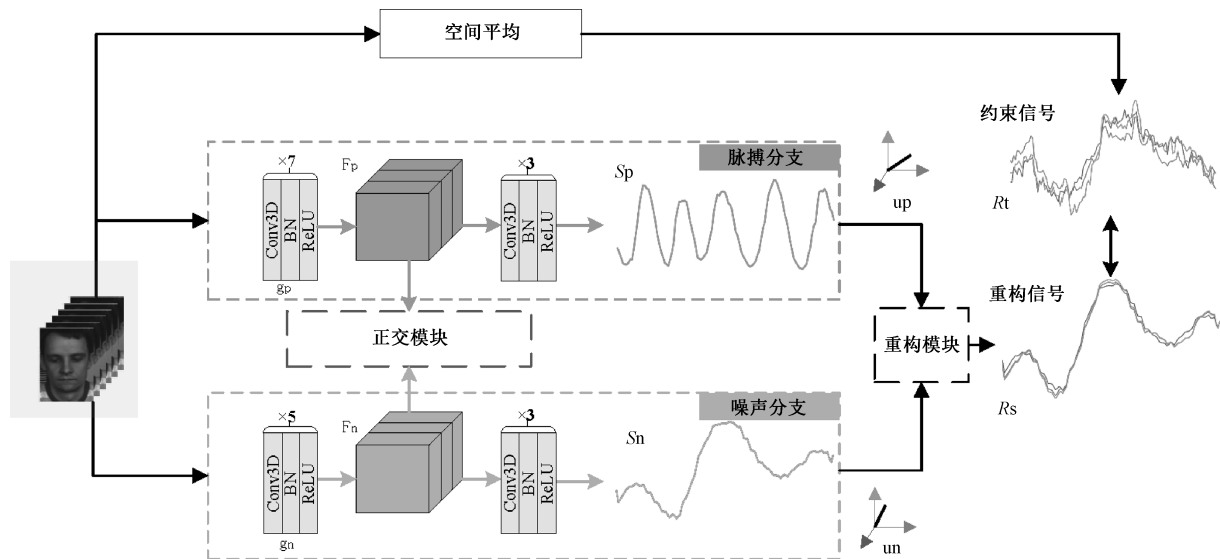


图1 PNRNet 总体框架示意图

Fig. 1 Overall framework of PNRNet

1.2 双分支脉搏—噪声建模

脉搏分支主要负责捕捉与心血管搏动同步的微弱颜色振荡。该分支通过多层 3D 卷积在空间维度聚合皮肤区域局部纹理与颜色变化, 在时间维度捕捉连续帧之间的周期性动态。卷积块通常由 3D 卷积、归一化、非线性激活和必要的时空下采样组成, 最终通过时间回归头输出长度为 T 的脉搏波形。

噪声分支与脉搏分支结构平行, 但其功能目标不同。该分支学习由头部运动、表情变化、光照波动、局部阴影、自动曝光、相机噪声等因素引起的非生理时序变化。若不显式设置噪声分支, 模型更容易将部分非生理变化混入脉搏表征, 从而产生噪声吸收现象; 引入噪声分支后, 模型具备了显式解释非生理变化的通道。

为了明确两条分支在网络中的输入、输出及其功能边界, 本文首先对脉搏分支和噪声分支的特征提取过程进行形式化描述。设输入面部视频片段为

F_p ; 噪声分支输出预测噪声信号 $\hat{s}_n$ 和噪声相关隐特征 F_n 。两分支均采用 3DCNN 作为前端时空特征提取器, 但参数彼此独立, 不共享前端特征。采用独立前端特征的原因在于, 脉搏信号通常表现为低幅值、周期性的局部颜色振荡, 而噪声更多表现为非周期、跨区域一致或突变式的颜色与亮度变化。独立建模有助于减少早期特征混合。

脉搏分支受参考 PPG 波形监督, 噪声分支则通过重构任务被迫解释输入 RGB 时序中不能由脉搏分量解释的残余扰动。两分支共同用于 RGB 观测重构, 使模型能够在脉搏波形预测与观测形成过程解释两个目标之间取得平衡。总体结构如图 1 所示。

$X \in \mathbb{R}^{B \times 3 \times T \times H \times W}$ 。脉搏分支和噪声分支分别通过相互独立的特征提取映射获得对应的隐特征:

$$F_p = g_p(X) \quad (1)$$

$$F_n = g_n(X) \quad (2)$$

式中: F_p 与 F_n 分别为脉搏特征和噪声特征; $g_p(\cdot)$ 和 $g_n(\cdot)$ 分别为脉搏分支和噪声分支中的特征提取映射, 其中 $g_p(\cdot)$ 通过 7 层卷积, $g_n(\cdot)$ 通过 5 层卷积, 每层卷积都通过 Conv3D、BN、ReLU 组成。

需要强调的是, 噪声分支不是简单地学习“错误”或“残差”, 而是学习一类与观测域扰动相对应的时间信号。该信号可与脉搏信号共同投影至 RGB 空间并重构 ROI 平均颜色序列, 因此具有明确的观测含义。通过保留此中间状态, 研究者能够观察模型是否把慢速光照漂移、运动突变或局部阴影等非生理因素分配给噪声路径。

1.3 基于正交约束的特征解耦

虽然两分支在结构上相互独立, 但如果缺少约束, 它

们仍可能在隐空间学习到重叠信息。例如,脉搏分支可能保留运动伪影,噪声分支也可能包含周期性脉搏成分。为此,本文在两分支的高层特征之间加入正交约束。设脉搏分支和噪声分支提取的高层特征分别为 $\mathbf{F}_p$ 和 $\mathbf{F}_n$ 。为避免空间位置差异对特征相关性度量造成干扰,本文首先对二者在空间维度上进行全局平均池化,将其压缩为通道-时间维度上的紧凑表征:

$$\tilde{\mathbf{F}}_p = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \mathbf{F}_p(h, w) \quad (3)$$

$$\tilde{\mathbf{F}}_n = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \mathbf{F}_n(h, w) \quad (4)$$

式中: H 和 W 分别表示特征图的高度和宽度,

$\tilde{\mathbf{F}}_p \in \mathbb{R}^{(3B) \times (THW)}$; $\tilde{\mathbf{F}}_n \in \mathbb{R}^{(3B) \times (THW)}$ 。记第 i 个样本的脉

搏特征向量和噪声特征向量分别为 $\tilde{\mathbf{F}}_p^{(i)}$ 和 $\tilde{\mathbf{F}}_n^{(i)}$ 。若两类特征中包含大量相同信息,则二者的余弦相似度会较高;反之,若两类特征趋于互补,其内积应接近于零。因此,本文将小批量样本内两类特征余弦相似度的绝对值作为正交约束损失:

$$\mathcal{L}_{\text{ortho}} = \frac{1}{B} \sum_{i=1}^B \left| \frac{\tilde{\mathbf{F}}_p^{(i)\text{T}} \times \tilde{\mathbf{F}}_n^{(i)}}{\|\tilde{\mathbf{F}}_p^{(i)}\|_2 \|\tilde{\mathbf{F}}_n^{(i)}\|_2} \right| \quad (5)$$

式中: $\mathcal{L}_{\text{ortho}}$ 表示正交约束损失; B 表示小批量样本数; i

表示批中第 i 个样本; $\tilde{\mathbf{F}}_p^{(i)}$ 和 $\tilde{\mathbf{F}}_n^{(i)}$ 分别表示对第 i 个样本的脉搏特征 $\mathbf{F}_p$ 和噪声特征 $\mathbf{F}_n$ 进行空间全局平均池化并向量化后的特征; $\|\cdot\|_2$ 表示 L2 范数。本文采用余弦相似度绝对值而非平方形式或铰链形式,主要基于两点考虑。首先,脉搏特征与噪声特征可能呈现正相关或负相关,只要二者存在较强线性相关性,均说明两分支表征仍有重叠。因此,对余弦相似度取绝对值可以同时惩罚正相关和负相关,使两类特征尽量不相关。其次,与平方惩罚相比,绝对值形式在相似度接近零时仍保持相对较高的约束信号,避免出现梯度消失,有助于持续抑制弱相关冗余;与铰链损失相比,绝对值形式不需要额外设置相关性容忍阈值,减少了超参数选择对训练结果的影响。

正交约束的作用是降低两分支表征相关性,引导脉搏相关特征和噪声相关特征趋于互补。与直接对输出波形做去相关不同,隐特征正交作用于更高维的时空表征空间,能够更早地影响特征学习方向。该约束并不意味着生理信号与干扰在真实世界中完全独立,而是在模型容量有限的条件下引导网络避免不必要的表征重叠。正交约束下两类特征的分离关系如图 2 所示。

需要说明的是,本文的正交约束作用于空间全局平均池化后的高层特征,主要约束通道-时间层面的整体相关性。该设计能够降低计算复杂度并避免局部空间错位

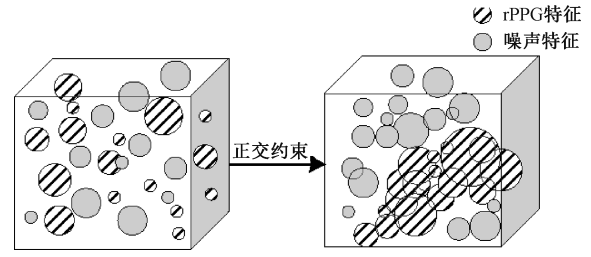


图 2 正交约束下脉搏特征与噪声特征解耦示意图

Fig. 2 Orthogonality-based feature disentanglement

对相关性度量造成过大影响,但也可能掩盖部分局部时空区域中的残余相关性。后续工作可考虑引入局部块级正交约束、注意力加权正交约束或多尺度相关性约束,以进一步增强局部时空解耦能力。

1.4 RGB 时序信号重构模块

RGB 重构模块是本文区别于普通双分支模型的关键部分。根据二色反射模型,皮肤像素颜色可看作光源强度、镜面反射、漫反射脉动成分和传感器噪声共同作用的结果。对 ROI 平均并进行均值去除、通道归一化后,静态肤色基线被弱化,剩余 RGB 时序变化可以近似表示为脉搏驱动变化与噪声驱动变化的叠加^[6-7]。

设预测脉搏信号为 $\hat{s}_p \in \mathbb{R}^{B \times T}$, 预测噪声信号为 $\hat{s}_n \in \mathbb{R}^{B \times T}$ 。为把一维时间信号映射回 RGB 观测域,引入可学习通道投影向量 $\mathbf{u}_p \in \mathbb{R}^{3 \times 1}$ 和 $\mathbf{u}_n \in \mathbb{R}^{3 \times 1}$, 分别表示脉搏分量和噪声分量在 RGB 通道上的相对响应方向;同时引入幅值系数 α_p 和 α_n 控制两类分量在重构中的强度。为提高模型约束,本文并未将 $\mathbf{u}_p$ 设置为完全自由的 RGB 投影方向。根据二色反射模型和 POS 等方法中的相关假设,脉搏相关的漫反射变化在 RGB 通道上的相对投影可近似约束为 $G > B > R$ 。因此,本文将该先验作为 $\mathbf{u}_p$ 的物理约束,使脉搏分量的 RGB 投影方向保持与皮肤漫反射脉搏响应一致的通道关系。与由脉搏血容量变化主导的生理信号方向不同,非生理扰动通常由多种复杂因素共同引起,这些噪声成分在 RGB 空间中的作用方向和强度并不固定,并且会随被试、场景和时间发生变化。因此,对 $\mathbf{u}_n$ 预设固定方向或施加显式方向约束难以准确刻画真实噪声分布,甚至可能限制模型对复杂扰动的表达能力。基于上述考虑,本文未对噪声方向 $\mathbf{u}_n$ 设置额外的方向约束,而是将其建模为可学习的非生理扰动方向,由模型根据输入 RGB 观测数据自适应估计。

由此可得到脉搏驱动 RGB 分量 $\mathbf{C}_p$ 和噪声驱动 RGB 分量 $\mathbf{C}_n$, 二者相加得到重构 RGB 时序信号 $\mathbf{R}_i$ 。监督信号 $\mathbf{R}_i$ 由面部 ROI 区域逐帧求平均得到原始 RGB 时间序列,再进行通道均值去除和归一化。该重构任务使网络在训练过程中同时利用脉搏与噪声分解释释真实视频观测,而不是仅优化单个输出波形与标签的一致性。该投

影与重构过程如图3所示。

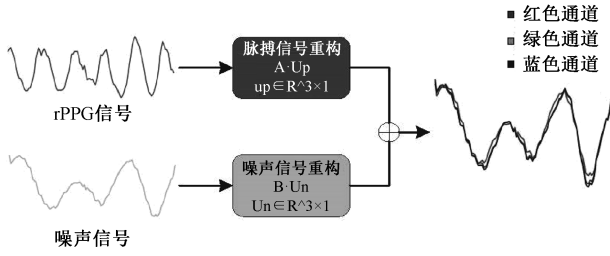


图3 RGB时序信号重构模块示意图

Fig. 3 RGB temporal signal reconstruction module

重构机制带来3方面优势:1)增加观测域监督,缓解单一PPG标签监督不足的问题;2)将神经网络输出与经典皮肤反射模型联系起来,提高物理可解释性;3)使噪声分支获得明确训练信号,避免噪声路径退化为无意义特征。

$$\mathbf{C}_p = \alpha_p \times \mathbf{u}_p \times \hat{s}_p \quad (6)$$

$$\mathbf{C}_n = \alpha_n \times \mathbf{u}_n \times \hat{s}_n \quad (7)$$

$$\mathbf{R}_s = \mathbf{C}_p + \mathbf{C}_n \quad (8)$$

$$\bar{\mathbf{X}} = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W \mathbf{X}(h, w) \quad (9)$$

$$\mathbf{R}_t = \mathbf{X} - \bar{\mathbf{X}} \quad (10)$$

$$\mathcal{L}_{\text{rec}} = 1 - \rho(\mathbf{R}_t, \mathbf{R}_s) \quad (11)$$

式中: $\mathcal{L}_{\text{rec}}$ 表示RGB重构损失; $\mathbf{C}_p$ 表示脉搏驱动RGB分量; $\mathbf{C}_n$ 表示噪声驱动RGB分量; $\hat{s}_p \in \mathbb{R}^{B \times T}$ 表示预测脉搏信号; $\hat{s}_n \in \mathbb{R}^{B \times T}$ 表示预测噪声信号; $\mathbf{u}_p \in \mathbb{R}^{3 \times 1}$ 表示脉搏分量在RGB三通道上的可学习投影方向; $\mathbf{u}_n \in \mathbb{R}^{3 \times 1}$ 表示噪声分量在RGB三通道上的可学习投影方向; α_p 和 α_n 为可学习幅值系数; $\mathbf{R}_s$ 表示重构得到的RGB时序信号; $\mathbf{R}_t$ 表示原始RGB时间序列。

1.5 损失函数与优化目标

本文采用联合目标函数优化PNRNet。总损失由脉搏波形监督损失、RGB重构损失、最大互相关损失、正交损失和频域分离损失组成。不同损失项从波形一致性、观测域一致性、时间错位容忍、特征解耦和频谱差异等角度共同约束模型。

脉搏监督损失采用负Pearson相关系数,强调预测波形与参考PPG在形状和趋势上的一致性。相比逐点MSE, Pearson损失对幅值尺度变化不敏感,更适合rPPG中归一化脉搏波形学习。RGB重构损失同样采用负Pearson相关,衡量重构RGB时间序列与ROI真实RGB时间序列之间的趋势一致性。

$$\mathcal{L}_{\text{pearson}} = 1 - \rho(\hat{s}_p, s) \quad (12)$$

式中: $\mathcal{L}_{\text{pearson}}$ 表示脉搏监督损失。

最大互相关损失用于处理面部rPPG与接触式参考

PPG之间可能存在的轻微相位差。由于面部脉搏波与接触式指夹或腕部PPG采集位置不同,血流传播路径和局部血流动力学差异可能导致峰值位置不完全一致。MCC损失在有限时间偏移范围内寻找最大相关值,使训练对小范围相位错位更加鲁棒。相位偏移与最大互相关对齐关系如图4所示。

$$MCC(\hat{s}_p, s) = \max_{\tau \in [-\Delta, \Delta]} \text{Corr}(\hat{s}_p(t), s(t - \tau)) \quad (13)$$

$$\mathcal{L}_{\text{mcc}} = 1 - MCC(\hat{s}_p, s) \quad (14)$$

式中: $\mathcal{L}_{\text{mcc}}$ 表示最大互相关损失; $MCC(\hat{s}_p, s)$ 表示预测脉搏信号与参考PPG信号之间的最大互相关; τ 表示时间偏移量; Δ 表示允许搜索的最大时间偏移范围; $\text{Corr}(\cdot, \cdot)$ 表示归一化互相关函数; $\max$ 表示在 $[-\Delta, \Delta]$ 范围内选择相关性最高的对齐结果。

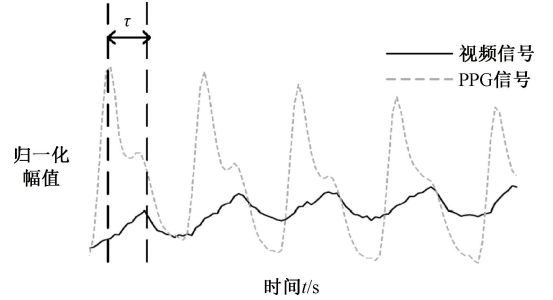


图4 最大互相关损失中的相位偏移示意图

Fig. 4 Phase shift in MCC supervision

频域分离损失用于鼓励脉搏信号和噪声信号在功率谱上形成差异。脉搏信号应在合理心率频段呈现明显主峰,而噪声信号可包含慢速光照漂移、高频传感器噪声或运动扰动。通过频域约束,可以降低噪声分支吸收真实脉搏周期成分的风险。

为计算频域差异,本文对脉搏分支输出信号和噪声分支输出信号分别去均值,并通过单边快速傅里叶变换获得功率谱密度。随后在0.6~4.0 Hz的生理心率频段内提取PSD,并对该频段内的PSD进行归一化。设归一化后的脉搏PSD和噪声PSD分别为 $P_{p(k)}$ 和 $P_{n(k)}$,目标频段内的频率bin数为 K ,则二者的频谱差异定义为:

$$d_{\text{PSD}} = (1/K) \times \sum_{k=1}^K (P_{p(k)} - P_{n(k)})^2 \quad (15)$$

频域分离损失通过增大 d_{PSD} 促使脉搏分支与噪声分支在频谱分布上保持差异,从而减少噪声分支对脉搏主频成分的吸收。

$$\mathcal{L}_{\text{psd}} = \max(0, m - d_{\text{PSD}}) \quad (16)$$

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{pearson}} + \lambda_2 \mathcal{L}_{\text{mcc}} + \lambda_3 \mathcal{L}_{\text{rec}} + \lambda_4 \mathcal{L}_{\text{ortho}} + \lambda_5 \mathcal{L}_{\text{psd}} \quad (17)$$

式中: $\mathcal{L}_{\text{total}}$ 表示总训练目标; $\mathcal{L}_{\text{psd}}$ 表示频域分离损失; λ_1 、 λ_2 、 λ_3 、 λ_4 和 λ_5 分别为各损失项的权重系数,用于平衡不同监督目标在联合优化中的贡献; m 表示预设间隔阈值;

$\max(0, m - d_{\text{psd}})$ 表示当频谱差异小于 m 时产生惩罚, 当差异达到或超过 m 时损失为 0, 本文方法 m 设置为 0.5。

2 实验设置

2.1 数据集

实验使用 iBVP、UBFC-rPPG 和 PURE 3 个公开数据集。iBVP 数据集^[25] 包含面部视频与同步生理信号, 具有一定光照和运动变化, 适合评估模型在复杂采集条件下的鲁棒性。UBFC-rPPG 数据集^[26] 采集场景相对稳定, 但仍包含自然面部运动和相机噪声, 是 rPPG 研究中常用的基准数据集。PURE 数据集^[27] 包含受试者在不同头部运动模式下的视频序列, 可用于检验模型在结构化运动干扰下的信号恢复能力。这 3 个数据集分别覆盖了相对稳定场景、复杂采集条件和结构化运动干扰, 因此可从不同角度评估 rPPG 方法的稳定性、运动鲁棒性和信号质量。

对于 RGB 重构监督, 逐帧计算 ROI 区域的 RGB 空间均值, 得到三通道时间序列。由于面向 rPPG 的人脸检测与跟踪质量会直接影响 ROI 区域的时序稳定性和后续脉搏信号恢复效果, 本文在获得 ROI 均值序列后进一步进行通道均值去除和归一化, 作为 RGB 重构目标 R_s 。

2.2 实现细节与评价指标

实验按照训练集、验证集和测试集 6 : 2 : 2 的比例划分。模型训练批大小设置为 4, 训练轮数设置为 10。优化过程中使用验证集监控模型性能, 最终在测试集上报告结果。所有对比方法均以相同指标进行评估, 以保证结果具有可比性。

评价指标包括信噪比 (signal-to-noise ratio, SNR)、平均绝对误差 (mean absolute error, MAE) 和均方根误差 (root mean square error, RMSE)。其中 SNR 越高表示恢复脉搏信号质量越好; MAE 和 RMSE 越低表示心率估计误差越小。RMSE 对大误差更加敏感, 因此可反映模型在困难样本上的稳定性。计算公式如下:

$$\text{SNR} = 10 \lg \frac{P_s}{P_n} \tag{18}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\widehat{HR}_i - HR_i| \tag{19}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\widehat{HR}_i - HR_i)^2} \tag{20}$$

式中: P_s 表示脉搏相关频率范围内的信号功率; P_n 表示非目标频率成分对应的噪声功率; N 表示测试样本数量; $\widehat{HR}_i$ 表示第 i 个样本的预测心率; HR_i 表示对应的真实心率。

对比方法包括传统信号处理方法 POS^[10]、

CHROM^[13] 以及深度学习方法 JAMSNet^[22]、PhysNet^[19] 和 GLISNet^[23]。此外, 本文设置跨数据集实验, 采用一个数据集训练、另一个数据集测试的方式, 分别考察 UBFC-rPPG 到 iBVP、PURE 到 UBFC-rPPG 以及 iBVP 到 PURE 的迁移性能。消融实验分别去除正交约束、重构模块、MCC 损失, 以分析各组件对最终性能的贡献。考虑到噪声分支是完整模型定义中的基础组成部分, 本文未设置“w/o Noise Branch”的消融项, 而重点比较正交约束、RGB 重构和 MCC 损失设计对性能的影响。

3 实验结果与分析

3.1 与主流方法的对比分析

表 1 ~ 3 为所提方法与 POS、CHROM、JAMSNet、PhysNet 和 GLISNet 在 iBVP、UBFC-rPPG 和 PURE 上的对比结果。整体来看, 本文方法在 3 个数据集上具有一定竞争力, 但不同数据集和不同评价指标上的优势并不完全一致。

表 1 iBVP 数据集上的定量比较

Table 1 Quantitative comparison on iBVP

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
POS ^[10]	1.00	8.71	15.05
CHROM ^[13]	1.00	8.46	14.85
JAMSNet ^[22]	3.89	4.53	8.94
PhysNet ^[19]	5.00	3.31	6.98
GLISNet ^[23]	3.15	4.17	9.47
本文	5.44	3.02	6.39

表 2 UBFC-rPPG 数据集上的定量比较

Table 2 Quantitative comparison on UBFC-rPPG

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
POS ^[10]	6.16	2.99	6.15
CHROM ^[13]	4.87	4.03	8.83
JAMSNet ^[22]	10.07	0.95	1.44
PhysNet ^[19]	9.66	1.33	1.82
GLISNet ^[23]	7.12	1.59	3.08
本文	10.61	1.18	1.64

表 3 PURE 数据集上的定量比较

Table 3 Quantitative comparison on PURE

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
POS ^[10]	4.71	2.87	6.35
CHROM ^[13]	3.25	3.38	9.20
JAMSNet ^[22]	6.27	1.46	3.75
PhysNet ^[19]	6.33	1.92	3.63
GLISNet ^[23]	5.51	1.77	4.30
本文	6.55	1.78	2.70

在 iBVP 数据集上,本文方法获得 5.44 的 SNR、3.02 的 MAE 和 6.39 的 RMSE,3 项指标均优于对比方法。iBVP 中存在相对复杂的采集条件,传统 POS 和 CHROM 由于依赖固定颜色投影假设,在噪声和运动条件下误差较大;深度方法虽然能够学习更复杂的表征,但若缺少显式噪声路径,仍可能把部分干扰混入脉搏特征。本文方法通过噪声分支吸收非生理扰动,并利用 RGB 重构约束两分支共同解释观测变化,因此在困难样本上 RMSE 更低。

在 UBFC-rPPG 数据集上,本文方法取得 10.61 的 SNR、1.18 的 MAE 和 1.64 的 RMSE。与 JAMSNet 相比,PNRNet 的 SNR 更高,说明其恢复信号中脉搏主频相对更加突出;但 JAMSNet 在 MAE 和 RMSE 上分别为 0.95 和 1.44,优于本文方法。因此,本文方法在该数据集上的优势主要体现在信号质量指标,而非所有误差指标。可能原因是 UBFC-rPPG 采集环境相对稳定,JAMSNet 的联合注意力和多尺度融合结构能够获得较低的心率估计误差;而 PNRNet 引入噪声分支和 RGB 重构约束后,更强调对观测形成过程和频谱质量的解释,因此在 SNR 上表现更突出。该结果也说明,不同模型在心率点估计误差与波形信号质量之间可能存在侧重点差异。

在 PURE 数据集上,本文方法的 MAE 为 1.78,RMSE 为 2.70,SNR 为 6.55。该结果表明,尽管部分方法在平均绝对误差上略有优势,但本文方法获得了最低 RMSE 和最高 SNR。由于 PURE 包含结构化头部运动,RMSE 和 SNR 的提升说明所提噪声分支能够在一定程度上解释运动引起的观测变化,同时保持脉搏成分的周期性。

表 4~6 为跨数据集泛化实验结果。结果显示,在 UBFC-rPPG→iBVP 设置下,本文方法取得 2.38 的 SNR、4.24 的 MAE 和 9.94 的 RMSE,3 项指标均优于 PhysNet 和 GLISNet;在 iBVP→PURE 设置下,本文方法的 SNR 为 -0.74,MAE 和 RMSE 分别为 6.15 和 7.36,也明显优于两种对比深度模型;在 PURE→UBFC-rPPG 设置下,本文方法的 MAE 和 RMSE 分别为 7.53 和 10.31,低于 PhysNet 和 GLISNet,SNR 也高于两者。上述结果表明,当训练集与测试集在采集设备、光照条件、运动模式和被试分布上存在差异时,显式的脉搏—噪声解耦与 RGB 观测域重构有助于降低模型对单一数据集特征分布的依赖,从而提升跨场景心率估计的稳定性。

表 4 UBFC-rPPG→iBVP 跨数据集上的定量比较

Table 4 Quantitative comparison on UBFC-rPPG→iBVP

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
PhysNet ^[19]	0.34	9.17	14.53
GLISNet ^[23]	-4.23	26.52	34.62
本文	2.38	4.24	9.94

表 5 iBVP→PURE 跨数据集上的定量比较

Table 5 Quantitative comparison on iBVP→PURE

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
PhysNet ^[19]	-2.87	21.84	25.04
GLISNet ^[23]	-3.74	15.66	19.82
本文	-0.74	6.15	7.36

表 6 PURE→UBFC-rPPG 跨数据集上的定量比较

Table 6 Quantitative comparison on PURE→UBFC-rPPG

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
PhysNet ^[19]	-4.79	16.62	20.85
GLISNet ^[23]	-2.24	10.73	16.51
本文	-1.14	7.53	10.31

由表 1~6 的对比结果可以看出,本文方法在 3 个数据集上均具有较强竞争力,但不同数据集和不同指标上的优势并不完全一致。PNRNet 在 iBVP 上 3 项指标均表现较好,在 UBFC-rPPG 上主要体现为 SNR 优势,在 PURE 上体现为较高 SNR 和较低 RMSE。上述结果说明,双分支解耦与 RGB 重构机制更有利于提升恢复波形的信号质量和困难样本下的稳定性。同时,跨数据集实验的误差整体仍高于同数据集测试,说明 rPPG 模型在跨设备、跨光源和跨运动模式条件下仍存在明显域偏移。后续可进一步结合域泛化训练、自监督预训练或测试时质量评估机制,提高模型在真实开放场景中的稳定性。

3.2 消融实验分析

表 7~9 为去除不同组件后的性能变化。与完整模型相比,去除正交约束、RGB 重构模块或 MCC 损失都会导致至少部分指标下降,说明各组件对最终性能均有积极作用。

表 7 iBVP 数据集上的消融实验

Table 7 Ablation results on iBVP

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
w/o Orth.	5.14	3.03	6.41
w/o Rec.	5.00	3.31	6.98
w/o MCC	5.06	3.35	7.00
本文	5.44	3.02	6.39

表 8 UBFC-rPPG 数据集上的消融实验

Table 8 Ablation results on UBFC-rPPG

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
w/o Orth.	9.94	1.35	2.05
w/o Rec.	9.92	1.35	2.04
w/o MCC	9.87	1.34	2.01
本文	10.61	1.18	1.64

表 9 PURE 数据集上的消融实验

Table 9 Ablation results on PURE

方法	SNR/dB ↑	MAE/bpm ↓	RMSE/bpm ↓
w/o Orth.	5.57	3.42	7.76
w/o Rec.	6.33	1.92	3.62
w/o MCC	5.62	3.67	7.56
本文	6.55	1.78	2.70

去除正交约束后, PURE 数据集上的 MAE 由 1.78 上升到 3.42, RMSE 由 2.70 上升到 7.76, 误差增幅较大。这表明在没有显式表征分离约束的情况下, 脉搏分支和噪声分支容易学习到重叠信息, 噪声路径可能吸收部分脉搏周期, 脉搏路径也可能保留运动或光照扰动, 最终影响波形质量。

去除重构模块后, 模型不再受到 ROI 级 RGB 时序观测的约束, 只依赖脉搏标签和其他正则项完成训练。此时 iBVP 上的 MAE 和 RMSE 均明显变差, 说明重构监督能够有效补充单一脉搏标签信息, 并帮助网络保持与真实视频观测的一致性。对于物理先验模型而言, 重构模

块不仅是辅助任务, 更是把脉搏—噪声分解与 RGB 成像机制连接起来的关键桥梁。

去除 MCC 损失后, 模型在多个数据集上的误差也出现上升, 尤其 PURE 上 MAE 和 RMSE 均明显变差。原因在于面部 rPPG 和接触式参考 PPG 之间可能存在自然相位差, 如果采用严格逐点对齐监督, 模型可能受到不必要的惩罚。MCC 损失通过在有限偏移范围内寻找最大相关, 使训练更加关注波形形态而非绝对峰值位置。

消融结果表明, 正交约束、RGB 重构模块和 MCC 损失均对性能提升具有积极作用。

3.3 波形分析

除定量指标外, rPPG 任务还需要关注预测波形与真实 PPG 信号在时域形态上的一致性。SNR、MAE 和 RMSE 能够从频域质量和心率估计误差角度评价模型性能, 但这些指标难以完全反映预测波形在主周期、峰谷位置、局部波动以及噪声抑制方面的表现。因此, 本文进一步在 iBVP 数据集上选取典型样本进行波形可视化分析, 如图 5 所示。

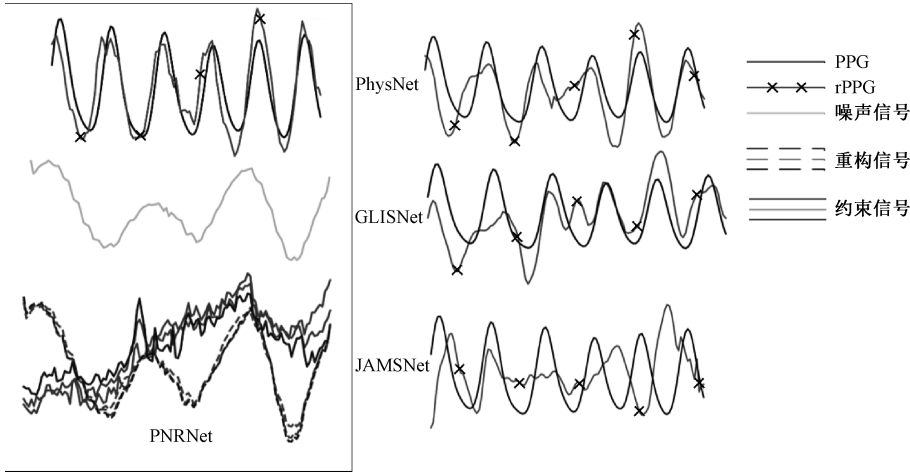


图 5 各方法预测 rPPG、真实 PPG 及噪声分支输出的波形对比示意图

Fig. 5 Waveform comparison of predicted rPPG signals, ground-truth PPG signals, and noise branch outputs across different methods

iBVP 数据集包含较明显的运动、光照变化和成像噪声等干扰, 面部 RGB 时序中的脉搏相关变化容易与非生理扰动发生混合。由图 5 可以看出, 传统或单分支深度学习学习方法在部分片段中虽然能够恢复出一定的周期性变化, 但预测波形与真实 PPG 之间仍存在不同程度的峰谷偏移、局部振荡异常或趋势不一致现象。例如, PhysNet、GLISNet 和 JAMNet 的预测 rPPG 在若干周期内能够跟随真实 PPG 的整体起伏, 但在局部峰值位置、谷值深度和波形平滑性方面仍存在偏差。这说明在噪声较强的样本中, 仅依赖端到端回归容易使模型将部分运动、光照或随机扰动混入脉搏表征, 进而影响恢复波形的稳定性。

势上与真实 PPG 保持了更好的一致性。图 5 中蓝色预测波形与黑色真实 PPG 在多个连续心动周期内具有较接近的峰谷结构, 局部异常振荡相对较少, 说明所提方法能够在复杂干扰条件下较稳定地提取脉搏相关成分。该现象与本文的双分支解耦设计一致: 脉搏分支主要学习与血容量变化相关的周期性信号, 噪声分支则用于吸收输入视频中由头部运动、光照波动、阴影变化和成像噪声引起的非生理成分。通过特征正交约束, 两类表征之间的冗余耦合被进一步降低, 从而减少噪声成分进入脉搏路径的可能性。

综合来看, 波形可视化结果进一步说明, 在 iBVP 这类噪声较多的数据集上, PNRNet 能够将部分非生理扰动

相比之下, PNRNet 的预测 rPPG 在主周期和整体趋

从脉搏表征中分离出来,使生成的 rPPG 信号在时域形态上更接近真实 PPG。该结果与前文定量实验中 iBVP 数据集上 SNR 提升、RMSE 降低的现象相互印证,说明脉搏—噪声双分支解耦与 RGB 时序重构机制有助于提高模型在复杂采集条件下的鲁棒性和波形恢复质量。

3.4 综合讨论

综合实验结果可见,PNRNet 的性能提升可能与显式脉搏—噪声建模及 RGB 观测域重构有关,而不仅仅来自网络复杂度的增加。传统深度模型通常依赖更强编码器或注意力模块提升回归精度,但内部中间状态难以解释。本文方法通过引入可观察的噪声路径,使模型能够把部分非生理变化从脉搏路径中分离出来,这对于真实系统中的异常检测、质量评估和临床可解释性均具有意义。

此外,RGB 重构模块使模型训练不再只依赖接触式 PPG 标签。实际数据采集中,参考 PPG 可能存在接触不稳、采样延迟或局部血流差异等问题。如果仅使用参考 PPG 逐点监督,模型可能受到标签噪声影响。重构监督则来自视频本身的 ROI 颜色变化,可在一定程度上补充标签监督,增强模型对观测形成过程的约束。

本文方法仍存在以下局限性。首先,噪声分支虽能建模非生理扰动,但解释粒度仍较粗,尚未进一步区分光照、运动、阴影和传感器噪声等具体来源。其次,本文的正交约束主要作用于全局池化后的高层特征,可能仍无法完全消除局部时空区域中的残余相关性。再次,RGB 重构模块中的通道投影方向为有效观测方向,尚未显式加入严格的皮肤反射物理约束。最后,虽然本文补充了跨数据集实验,但跨设备、跨肤色、跨光源和跨压缩质量条件下的泛化能力仍需要更系统的验证。

未来工作可从 3 方面展开:1) 将噪声分支进一步细分为运动、光照和成像噪声子分支,提高可解释粒度;2) 引入自监督预训练和域泛化策略,增强跨数据集迁移能力;3) 结合质量评估机制,根据噪声分支强度或频谱结构自动判断当前 rPPG 输出是否可信。

4 结 论

本文提出 PNRNet,一种面向鲁棒远程光电容积描记的物理先验引导双分支时空解耦与重构框架。该方法以相互独立的脉搏分支和噪声分支分别建模生理脉搏变化与非生理扰动变化,并在隐特征空间引入正交约束以增强表征分离。受二色反射模型启发,本文进一步设计 RGB 时序信号重构模块,将预测脉搏与噪声信号投影回 RGB 观测域,使网络训练同时满足脉搏波形监督和视频观测域一致性约束。该机制提高了模型的物理可解释性,也使中间噪声分量具备可审查价值。在 iBVP、UBFC-rPPG 和 PURE 数据集上的实验结果表明,所提方法在

SNR、MAE 和 RMSE 指标上具有较强竞争力,并在波形稳定性和信号质量方面表现较好。消融实验进一步验证了正交约束、重构模块、MCC 损失的有效性。未来将结合更细粒度噪声建模、跨域预训练和真实场景质量控制机制,进一步提升非接触式生理信号恢复的鲁棒性与可解释性。

参考文献

- [1] 牛雪松, 韩琥, 山世光. 基于 rPPG 的生理指标测量方法综述[J]. 中国图象图形学报, 2020, 25(11): 2321-2336.
- Niu Xuesong, Han Hu, Shan Shiguang. Survey on rPPG-based physiological measurement methods[J]. Journal of Image and Graphics, 2020, 25(11): 2321-2336.
- [2] Zhao Changchen, Zhang Shunhao, Cao Pengcheng, et al. Pruning remote photoplethysmography networks using weight-gradient joint criterion[J]. Expert Systems with Applications, 2025, 282: 127623.
- [3] Zheng Xiujuan, Zou Binghang, Zhang Chang, et al. Remote blood pressure estimation using BVP signal features from facial videos[J]. Pattern Recognition Letters, 2025, 193: 122-127.
- [4] Liu Bobiang, Zheng Xiujuan, Huang Qiu, et al. Remote respiratory rate measurement from torso video in challenging scenarios[J]. Optics Express, 2026, 34(5): 8579-8597.
- [5] Verkrusse W, Svaasand L O, Nelson J S. Remote plethysmographic imaging using ambient light[J]. Optics Express, 2008, 16(26): 21434-21445.
- [6] Da Costa G. Optical remote sensing of heartbeats[J]. Optics Communications, 1995, 117(5-6): 395-398.
- [7] Shafer S A. Using color to separate reflection components[J]. Color Research & Application, 1985, 10(4): 210-218.
- [8] Liu Boxiang, Zheng Xiujuan, Wu Y I. Remote heart rate estimation in intense interference scenarios: A white-box framework[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-14.
- [9] Poh M Z, McDuff D J, Picard R W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation[J]. Optics Express, 2010, 18(10): 10762-10774.
- [10] De Haan G, Jeanne V. Robust pulse rate from chrominance-based rPPG[J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2878-2886.
- [11] De Haan G, Van Leest A. Improved motion robustness of remote-PPG by using the blood volume pulse signature[J]. Physiological Measurement, 2014, 35(9): 1913-1926.
- [12] Wang Wenjin, Stuijk S, De Haan G. A novel algorithm for

- remote photoplethysmography: Spatial subspace rotation[J]. IEEE Transactions on Biomedical Engineering, 2016, 63(9): 1974-1984.
- [13] Wang Wenjin, Den Brinker A C, Stuijk S, et al. Algorithmic principles of remote PPG [J]. IEEE Transactions on Biomedical Engineering, 2017, 64(7): 1479-1491.
- [14] 刘祎, 欧阳健飞, 闫勇刚. 基于普通摄像头的心率测量方法研究[J]. 计算机工程与应用, 2016, 52(7): 210-214.
- Liu Yi, Ouyang Jianfei, Yan Yonggang. Research on heart rate measurement method based on ordinary camera[J]. Computer Engineering and Applications, 2016, 52(7): 210-214.
- [15] 戚刚, 杨学志, 吴秀, 等. 非合作面部晃动情况下的心率检测[J]. 中国图象图形学报, 2017, 22(1): 126-136.
- Qi Gang, Yang Xuezhi, Wu Xiu, et al. Heart rate detection under non-cooperative facial shaking [J]. Journal of Image and Graphics, 2017, 22(1): 126-136.
- [16] 赵彦涛, 付美玲, 王斌, 等. 基于主成分分析和重叠直方图统计的视频信号心率测量[J]. 中国生物医学工程学报, 2016, 35(3): 284-291.
- Zhao Yantao, Fu Meiling, Wang Bin, et al. Video-based heart rate measurement using principal component analysis and overlapped histogram statistics[J]. Chinese Journal of Biomedical Engineering, 2016, 35(3): 284-291.
- [17] 赵昶辰, 梅培义, 冯远静. 面向远程光体积描记的人脸检测与跟踪[J]. 中国图象图形学报, 2020, 25(11): 2439-2450.
- Zhao Changchen, Mei Peiyi, Feng Yuanjing. Face detection and tracking algorithm for remote photoplethysmography [J]. Journal of Image and Graphics, 2020, 25(11): 2439-2450.
- [18] Chen Weixuan, McDuff D. DeepPhys: Video-based physiological measurement using convolutional attention networks[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2018: 347-364.
- [19] Yu Zitong, Li Xiaobai, Zhao Guoying. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks [C]//Proceedings of the British Machine Vision Conference. Cardiff: BMVC, 2019.
- [20] Das A, Lu Hao, Han Hu, et al. BVPNet: Video-to-BVP signal prediction for remote heart rate estimation [C]//Proceedings of the 16th IEEE International Conference on Automatic Face and Gesture Recognition. IEEE, 2021: 1-8.
- [21] 赵昶辰, 居峰, 冯远静. 多视角神经网络非接触式脉搏信号提取[J]. 中国图象图形学报, 2020, 25(11): 2428-2438.
- Zhao Changchen, Ju Feng, Feng Yuanjing. Noncontact pulse signal extraction based on multiview neural network [J]. Journal of Image and Graphics, 2020, 25(11): 2428-2438.
- [22] Zhao Changchen, Wang Hongsheng, Chen Huiling, et al. JAMSNet: A remote pulse extraction network based on joint attention and multi-scale fusion [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(6): 2783-2797.
- [23] Zhao Changchen, Zhou Menghao, Zhao Zheng, et al. Learning spatio-temporal pulse representation with global-local interaction and supervision for remote prediction of heart rate [J]. IEEE Journal of Biomedical and Health Informatics, 2024, 28(2): 609-620.
- [24] Zhao Changchen, Cao Pengcheng, Hu Meng, et al. WTC3D: An efficient neural network for noncontact pulse acquisition in Internet of Medical Things [J]. IEEE Transactions on Industrial Informatics, 2025, 21(2): 1547-1556.
- [25] Joshi J N, Cho Y. iBVP dataset: RGB-thermal rPPG dataset with high resolution signal quality labels [J]. Electronics, 2024, 13(7): 1334.
- [26] McDuff D, Gontarek S, Picard R W. Remote measurement of cognitive stress via heart rate variability [C]//Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2014: 2957-2960.
- [27] Stricker R, Müller S, Gross H M. Non-contact video-based pulse rate measurement on a mobile service robot [C]//Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication. IEEE, 2014: 1056-1062.

作者简介



张建设, 1997 年于济南大学获得学士学位, 2004 年于浙江大学获得硕士学位, 2008 年于浙江大学获得博士学位, 现为杭州电子科技大学计算机学院教授, 主要研究方向为人工智能、类脑计算及脑-机接口技术。

E-mail: jhzhang@hdu.edu.cn

Zhang Jianhai received his B. Sc. degree from University of Jinan in 1997, M. Sc. degree from Zhejiang University in 2004, and Ph. D. degree from Zhejiang University in 2008. He is currently a professor with the School of Computer Science, Hangzhou Dianzi University. His main research interests include

artificial intelligence, brain-inspired computing, and brain-computer interface technology.



姜丰, 2024 年于浙江理工大学获得学士学位, 现为杭州电子科技大学硕士研究生, 主要研究方向为视觉非接触式生理信号监测、智慧医疗。

E-mail: 242050209@hdu.edu.cn

Jiang Feng received his B. Sc. degree from Zhejiang Sci-Tech University in 2024. He is currently a M. Sc. candidate at Hangzhou Dianzi University. His main research interests include visual non-contact physiological signal monitoring and smart healthcare.



程世超, 2013 年于河南师范大学获得学士学位, 2019 年于大连理工大学获得博士学位, 现为杭州电子科技大学计算机学院副教授, 主要研究方向为图像处理、可学习优化方法、计算机视觉、脑机接口。

E-mail: sccheng@hdu.edu.cn

Cheng Shichao received her B. Sc. degree from Henan Normal University in 2013 and Ph. D. degree from Dalian University of Technology in 2019. She is currently an associate

professor with the School of Computer Science, Hangzhou Dianzi University. Her main research interests include image processing, learnable optimization methods, computer vision, and brain-computer interfaces.



赵昶辰 (通信作者), 2011 年于山东科技大学获得学士学位, 2018 年于台湾中兴大学、北京航空航天大学获得博士学位, 现为杭州电子科技大学计算机学院副教授, 主要研究方向为视觉非接触式生理信号监测、心血管疾病风险预测、下肢康复机器人、智慧医疗。

E-mail: cczhao@hdu.edu.cn

Zhao Changchen (Corresponding author) received his B. Sc. degree from Shandong University of Science and Technology in 2011, and Ph. D. degree from National Chung Hsing University and Beihang University in 2018. He is currently an associate professor with the School of Computer Science, Hangzhou Dianzi University. His main research interests include visual contactless physiological signal monitoring, cardiovascular disease risk prediction, lower-limb rehabilitation robots, and smart healthcare.