

DOI: 10.13382/j.jemi.B2609266

# 融合深度可分离卷积与双向 Mamba 的 rPPG 提取<sup>\*</sup>

朱海军<sup>1</sup> 李文洋<sup>1</sup> 徐礼胜<sup>2</sup> 刘秀玲<sup>1</sup> 赵昶辰<sup>3</sup> 杜 硕<sup>1</sup>

(1. 河北大学电子信息工程学院 保定 071000; 2. 东北大学信息科学与工程学院 沈阳 110004;  
3. 杭州电子科技大学计算机学院 杭州 310018)

**摘 要:**无创非接触式生理参数测量是健康监测领域的重要研究方向,远程光电容积描记技术是该领域主流技术。针对现有 rPPG 算法难以同时兼顾检测精度、轻量化部署与长时序建模效率低的问题,提出了 rPPG-Mamba 模型。模型采用深度可分离卷积压缩模型开销,利用双向 Mamba 强化长时序特征建模,并引入时间差分卷积、3D 通道注意力提升复杂场景特征提取能力,实现人脸视频中 rPPG 信号精准提取。在 PURE、UBFC-rPPG、MMPD 数据集开展对比实验,在数据集内评估中,模型在 PURE 数据集上平均绝对误差达 0.18、均方根误差达 0.31,在 UBFC-rPPG 数据集上平均绝对误差达 0.43、均方根误差达 0.68,均取得最优性能;在跨数据集实验中,验证了模型良好的泛化性,在复杂干扰数据集 MMPD 中性能仅略低于 PhysMamba。模型参数量较 EfficientPhys 降低了约 89.9%,运算量较 DeepPhys 降低了 61.5%。实验表明,rPPG-Mamba 可兼顾精度与轻量化能力,具备实际应用前景,可为非接触远程生理测量提供技术参考。

**关键词:**深度可分离卷积;双向 Mamba;远程光电容积描记术;3D 时空特征提取

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

## Remote photoplethysmography extraction combining depthwise separable convolution and bidirectional Mamba

Zhu Haijun<sup>1</sup> Li Wenyang<sup>1</sup> Xu Lisheng<sup>2</sup> Liu Xiuling<sup>1</sup> Zhao Changchen<sup>3</sup> Du Shuo<sup>1</sup>

(1. School of Electronic Information Engineering, Hebei University, Baoding 071000, China;  
2. School of Information Science and Engineering, Northeastern University, Shenyang 110004, China;  
3. School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China)

**Abstract:** Non-invasive and contactless physiological parameter measurement is a vital research direction in health monitoring, and remote photoplethysmography (rPPG) has become a mainstream technique in this field. Existing rPPG algorithms struggle to balance detection accuracy and lightweight deployment, and suffer from low efficiency in modeling long-range temporal dependencies. To solve these problems, this paper proposes an rPPG-Mamba model. The model adopts depthwise separable convolutions to reduce parameters and computational overhead, and leverages bidirectional Mamba to perform efficient long-range temporal modeling. Temporal difference convolution and 3D channel attention are further integrated to improve feature extraction capability in complex scenarios, enabling accurate rPPG signal extraction from facial videos. Extensive experiments were conducted on three public datasets: PURE, UBFC-rPPG and MMPD. In intra-dataset evaluations, the proposed model achieved a mean absolute error (MAE) of 0.18 and a root mean square error (RMSE) of 0.31 on the PURE dataset, as well as an MAE of 0.43 and an RMSE of 0.68 on the UBFC-rPPG dataset, outperforming all comparison methods. Cross-dataset tests also verified its satisfactory generalization ability. On the MMPD dataset with severe illumination and motion interference, its performance was only slightly worse than that of PhysMamba. Compared with EfficientPhys, the proposed model reduced the parameter count by approximately 89.9%, while its computational complexity was 61.5% lower than that of DeepPhys. The experimental results demonstrate that rPPG-Mamba achieves an excellent trade-off between detection accuracy and lightweight design, and possesses great application potential for non-contact remote physiological measurement.

**Keywords:**depthwise separable convolution; bidirectional Mamba; remote photoplethysmography; 3D spatiotemporal feature extraction

收稿日期:2026-04-14 Received Date: 2026-04-14

<sup>\*</sup> 基金项目:河北省自然科学基金(H2025201053, F2025201070)、国家自然科学基金(52207251, 62273082, 62573171)项目资助

## 0 引言

生理参数(如心率、血氧饱和度、呼吸频率等)的实时、精准测量是医疗健康监测、人机交互、运动健康管理等领域的基础,其测量精度与便捷性直接影响健康评估的科学性和应用落地的可行性<sup>[1]</sup>。传统生理测量方法依赖接触式传感设备(如心电图机、脉搏血氧仪等),需与人体皮肤直接接触才能获得生理信号,存在操作繁琐、佩戴不适、无法实现长期连续监测等局限,难以满足居家健康监测、远程医疗、特殊场景(如新生儿监护、重症患者无创监测)下的应用需求<sup>[2]</sup>。

随着计算机视觉、数字信号处理与机器学习技术的快速发展,无创、非接触式生理测量技术成为研究热点,其中远程光电容积描记术(remote photoplethysmography, rPPG)凭借其操作便捷、无接触、低成本、可实现长期连续监测等优势,逐渐成为非接触生理测量领域的重要研究方向<sup>[3]</sup>。rPPG 技术基于光体积描记原理,利用普通摄像头捕捉人体面部、指尖等部位的视频图像,通过提取图像中因人体心血管搏动导致的皮肤颜色微小变化,推导出心率、血氧饱和度等关键生理参数,无需任何接触式传感设备,可实现自然、无感的生理监测<sup>[4]</sup>。

rPPG 技术作为非接触式生理参数测量的核心方法,在早期,rPPG 研究以手工特征设计与数字信号处理为核心,采用多阶段操作完成信号提取与参数估计,核心思路是通过颜色空间变换、盲源分离等手段从视频中分离生理信号与噪声<sup>[5]</sup>。Verkruysse 等<sup>[6]</sup>率先从 RGB 视频绿色通道中提取原始 rPPG 信号,验证了光电容积描记原理在非接触测量中的可行性;Poh 等<sup>[7]</sup>提出经典的 CHROM 算法,通过对 RGB 肤色通道进行线性归一化变换抑制光照干扰,成为后续传统方法的基础;De Haan 团队<sup>[8-9]</sup>相继提出 POS、2SR 等算法,通过构建颜色子空间变换模型克服运动伪影,提升了复杂场景下的信号质量。此外,部分研究引入独立成分分析、主成分分析等盲源分离技术,以及经验模态分解、小波滤波等信号处理方法,对提取的原始信号进行去噪与增强,进一步提升信号质量<sup>[10-12]</sup>。

随着深度学习技术的发展,卷积神经网络被引入到 rPPG 信号提取中,替代手工特征实现端到端的特征学习,有效突破了传统方法的缺陷。Chen 等<sup>[13]</sup>提出 DeepPhys 模型,设计双流 3D 卷积网络结合注意力机制,自适应聚焦面部生理信号敏感区域,实现了光照与运动干扰下的 rPPG 信号提取;Yu 等<sup>[14]</sup>提出 PhysNet 模型,构建时空分离的 3D 卷积结构,强化对视频帧间时序特征的捕捉能力,提升了长序列视频中信号提取的准确性;Niu 等<sup>[15]</sup>提出 RhythmNet 模型,通过卷积神经网络生成面部多区域时空特征图,结合循环神经网络建模时序依赖,实

现心率的精准回归。此外,TS-CAN、AutoHR、PhysFormer 等模型进一步优化了卷积结构与注意力机制,AutoHR 还引入神经架构搜索挖掘最优骨干网络,并结合时间差分卷积(temporal difference convolution, TDC)增强时序动态特征,大幅提升了信号提取的精度<sup>[16-18]</sup>。研究人员又提出直接从面部视频回归心率等生理参数的端到端 rPPG 方法,摒弃独立的 rPPG 信号提取过程。Bousefsaf 等<sup>[19]</sup>首次提出基于 3D 卷积的心率预测模型,通过仿真数据训练实现面部视频到心率的直接映射,验证了方法的可行性;Tulyakov 等<sup>[20]</sup>设计融合 3D 卷积与长短期记忆的时空网络,直接从 2 s 短时间窗口视频中估计心率,兼顾了瞬时心率检测效果与实时监测性能。然而,现有基于深度学习的 rPPG 方法仍存在诸多不足:一方面,部分模型结构复杂、计算量大,难以适配移动端、嵌入式等资源受限设备,限制了其在实时监测场景中的应用;另一方面,在复杂干扰场景下,现有方法的测量精度仍有较大提升空间。

3D 深度可分离卷积可在保留特征提取能力的前提下,有效降低模型参数量与计算开销,是 rPPG 轻量化网络设计的核心技术。Ouzar 等<sup>[21]</sup>提出 X-iPPGNet 模型,借鉴 Xception 网络架构,将 3D 深度可分离卷积引入到 rPPG 提取任务,通过颜色通道解耦分别学习各通道的光电容积信息,在降低计算量的同时提升了特征表达的丰富性,成为轻量化 rPPG 模型的代表。除 X-iPPGNet 外,EfficientPhys 等模型也将深度可分离卷积与注意力机制结合,构建轻量化的 rPPG 特征提取网络,实现了精度与效率的初步平衡<sup>[22-23]</sup>;Ma 等<sup>[24]</sup>将深度可分离卷积与残差连接融合,缓解了深层网络的梯度消失问题,进一步提升了模型的特征学习能力。尽管 3D 深度可分离卷积在轻量化上优势显著,但该技术 in rPPG 任务中的应用仍存在挑战:单独使用时,特征提取的表征能力有限,会使精度降低,且在时序维度的特征建模上存在不足,难以充分捕捉 rPPG 信号的长距离时序依赖。

近年来,以 Mamba 为代表的状态空间模型凭借线性时间复杂度与优异的长序列建模能力,成为时序信号处理领域的研究热点,也为 rPPG 长时序建模提供了新的技术路径。Gu 等<sup>[25]</sup>提出的基础 Mamba 模型,摒弃了 Transformer 的高计算开销,通过流式状态空间架构实现长序列的高效处理;在此基础上,PhysMamba 将 Mamba 引入到 rPPG 信号提取的任务,初步验证了状态空间模型在该领域的应用潜力,在长时序视频中展现出优于 Transformer 的建模效率<sup>[26]</sup>。针对单向 Mamba 在 rPPG 时序建模中存在的信息遗忘与位置偏差问题,Qian 等<sup>[27]</sup>进一步提出双向 Mamba 模型,引入前后双向扫描策略。该模型继承 Mamba 线性复杂度的特点,可完整捕获 rPPG 的上下文时序关联,弥补单向结构的缺陷,同步提升长视

频场景下的建模效率与信号提取精度,为 rPPG 长时序建模与远程生理参数监测提供了更具竞争力的技术路径。

针对现有 rPPG 技术存在的精度与轻量化难以兼顾、长时序依赖建模效率不足两大问题,本文对研究方法进行创新。首先,针对传统 3D 卷积模型参数量大、计算复杂度高的问题,将 3D 深度可分离卷积引入 rPPG 特征提取流程,在保证局部时空特征提取能力的同时,大幅降低模型参数量与运算开销,为设备端部署提供可行性<sup>[28]</sup>;其次,针对 Transformer 类模型长序列建模计算成本高、单向 Mamba 存在时序信息遗忘与位置偏差的问题,引入双向 Mamba 模块,通过前向-后向双向扫描机制,在保持线性时间复杂度的前提下,实现 rPPG 信号长距离时序依赖的精准捕捉,提升长时序建模效率与精度<sup>[29]</sup>;同时,为进一步增强复杂场景下的特征提取能力,在特征提取模块中嵌入 TDC 与 3D 通道注意力(channel attention, CA),前者放大皮肤血流产生的细微时序变化、抑制静态噪声,后者自适应强化有效特征通道、抑制冗余信息,实现复杂干扰下的信号鲁棒提取<sup>[30]</sup>。

基于上述思路,本文设计融合深度可分离卷积与双向 Mamba 的 rPPG-Mamba 双分支模型。结合深度可分离卷积的轻量化特征提取优势与双向 Mamba 的高效长序列建模能力,将 TDC 与 CA 加入到 3D 可分离卷积与双向 Mamba 中构成局部时空特征模块(local spatiotemporal features, LSF)和长时序依赖模块(long-term sequence dependency, LTD),以 3 个 LSF 和 LTD 模块构建空间、时序双分支架构,通过横向连接实现双分支特征的维度对齐与互补融合,最终实现面部视频中 rPPG 信号的高效、精准提取。

## 1 方法

### 1.1 TDC

TDC 作为时序特征增强单元,嵌入 3D 可分离卷积与双向 Mamba 中的特征提取流程,突出面部皮肤因血液流动产生的细微颜色变化。如图 1 所示,TDC 通过显式建模相邻帧间的时序差分信息,放大面部皮肤因血液流动产生的细微颜色变化,抑制静态背景与噪声干扰,增强 rPPG 信号的细粒度动态特征表达<sup>[31]</sup>,最终输出时序差分增强特征  $F_{tdc}$ 。TDC 作为 rPPG-Mamba 的基础时序增强单元,为后续时空特征提取提供高质量时序差分特征。其表达式如下:

$$F_{tdc} = \sum_{i=-k}^k w_i \times X_{t+i} - \theta \times X_t \quad (1)$$

式中:  $X_t$  为当前时刻  $t$  的输入特征图;  $X_{t+i}$  为当前时刻  $t$  相邻  $i$  帧的特征图;  $w_i$  为时序邻域的可学习权重系数;  $\theta \in [0,1]$  为差分强度调节因子;  $F_{tdc}$  为时序差分增强

后的输出特征图。

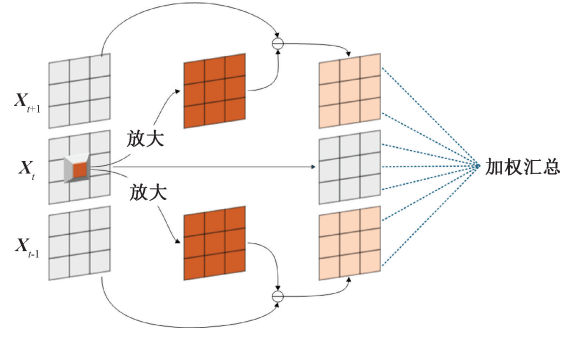


图1 TDC 模块结构

Fig. 1 TDC module structure diagram

### 1.2 3D 深度可分离卷积

3D 深度可分离卷积是 LSF 模块的核心特征提取单元,将标准 3D 卷积解耦为深度卷积与逐点卷积两个独立操作<sup>[32]</sup>,如图 2 所示。先进行深度卷积运算,对输入特征图  $X$  的每个通道独立应用 3D 卷积核  $K_d$ ,仅捕捉单个通道内的局部时空特征,输出深度卷积特征  $F_{dus}$ 。再进行逐点卷积运算,用  $1 \times 1 \times 1$  的卷积核  $K_p$  对深度卷积输出  $F_{dus}$  进行通道融合,将单通道局部特征整合为多通道全局特征,最终输出空间分支的局部时空特征  $F_{sep}$ 。3D 深度可分离卷积在保持局部时空特征提取能力的前提下,可以有效减少卷积核参数量与计算量,是实现 rPPG-Mamba 模型轻量化的关键。其表达式如下:

$$F_{dus} = \text{DepthwiseConv3d}(X, K_d, S, P) \quad (2)$$

$$F_{sep} = \text{PointwiseConv3d}(F_{dus}, K_p) \quad (3)$$

式中:  $X$  为输入时空特征图;  $K_d$  为深度卷积核;  $K_p$  为逐点卷积核;  $S, P$  分别为卷积操作的步长与填充大小;  $F_{dus}, F_{sep}$  分别为深度卷积、逐点卷积的输出特征图。

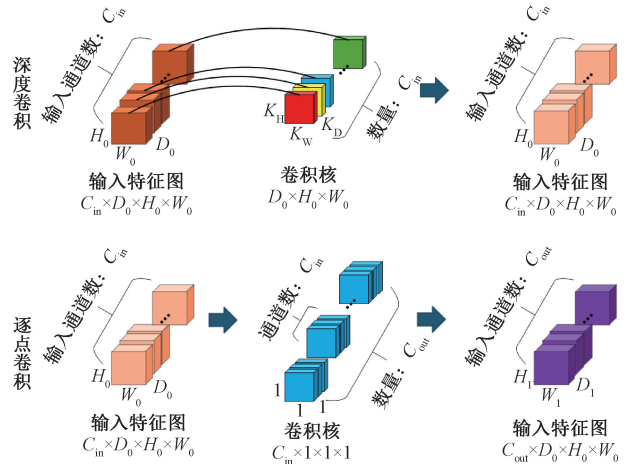


图2 深度卷积与逐点卷积结构

Fig. 2 Diagram of depthwise and pointwise convolution structures

### 1.3 双向 Mamba

双向 Mamba 是 LTD 模块的核心长时序建模单元,如图 3 所示,双向 Mamba 在基础 Mamba 的线性时间复杂度长序列建模基础上,引入前向-后向双向扫描机制,再进行双向特征融合与残差连接,避免深层网络梯度消失,解决单向 Mamba 的时序信息遗忘与位置偏差问题<sup>[33]</sup>,最终输出包含长时序依赖的特征  $F_{mamba \odot}$ 。

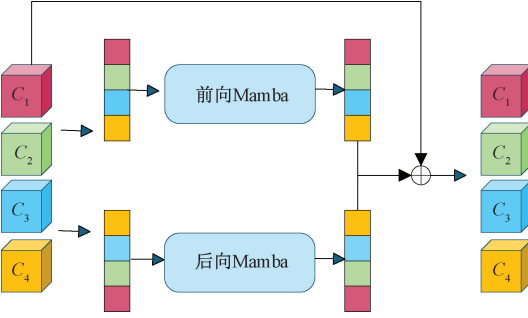


图 3 双向 Mamba 结构

Fig. 3 Bidirectional Mamba structure diagram

双向 Mamba 可同时捕捉 rPPG 信号的前向时序关联与后向上下文信息,精准建模长距离时空依赖,挖掘 rPPG 信号的周期性规律。双向 Mamba 是 rPPG-Mamba 长时序建模的核心模块,解决 rPPG 信号长依赖建模难的问题。其表达式如下:

$$h_t^f = A_t^f \odot h_{t-1}^f + B_t^f \odot U_t \quad (4)$$

$$h_t^b = A_t^b \odot h_{t+1}^b + B_t^b \odot U_t \quad (5)$$

$$F_{mamba} = \text{Linear}(h_t^f + h_t^b) + X \quad (6)$$

式中:  $h_t^f$ 、 $h_t^b$  分别为  $t$  时刻前向、后向传播的隐藏状态;  $A_t^f$ 、 $A_t^b$  分别为前向、后向的状态转移矩阵且不共享网络参数;  $B_t^f$ 、 $B_t^b$  分别为前向、后向的输入投影矩阵,  $B_t^f$ 、 $B_t^b$  参数同样相互独立;  $U_t$  为  $t$  时刻输入特征经线性层投影后的结果;  $X$  为 LTD 模块的原始输入特征;  $F_{mamba}$  为 LTD 模块的输出特征图。双向分支均沿用基础 Mamba 的选择性状态筛选机制,在各自扫描方向内依靠选择性参数动态筛选有效时序信息,双向架构仅改变特征遍历顺序,不破坏原生 Mamba 选择性筛选逻辑,二者架构完全兼容。

### 1.4 3D 通道注意力

CA 模块嵌入双分支特征提取流程的末端通过自适应学习通道维度的权重分布,强化对 rPPG 信号有贡献的有效特征通道,抑制无意义的冗余特征通道,提升特征表达的有效性<sup>[34]</sup>。如图 4 所示,CA 模块先进行全局时空平均池化,再进行通道权重学习,将池化结果  $z_c$  输入到两层全连接层的多层感知机进行非线性变换,再通过 Sigmoid 激活函数  $\sigma$  将输出映射至  $[0, 1]$  区间,得到第  $c$  个通道的权重系数  $s_c$ , 权重越大表示该通道对 rPPG 信

号提取的贡献越大。最后实现通道加权,强化高权重有效通道的特征,抑制低权重冗余通道,最终输出优化后的特征  $F_{ca}$ 。CA 模块是 rPPG-Mamba 双分支特征提取的关键优化单元,为空间、时序分支的特征输出做精细化筛选,进一步提升模型对 rPPG 有效信号的捕捉能力。其表达式如下:

$$z_c = \frac{1}{T \times H \times W} \sum_{t=1}^T \sum_{i=1}^H \sum_{j=1}^W F_{c,t,i,j} \quad (7)$$

$$s_c = \sigma(\text{MLP}(z_c)) \quad (8)$$

$$F_{ca} = s_c \odot F \quad (9)$$

式中:  $F_{c,t,i,j}$  为输入特征图  $F$  在第  $c$  个通道、第  $t$  个时间帧、空间位置  $(i, j)$  的像素值;  $z_c$  为第  $c$  个通道的全局时空平均池化结果; MLP 为多层感知机;  $\sigma$  为 Sigmoid 激活函数;  $s_c$  为每个通道的权重系数;  $\odot$  为逐元素相乘操作;  $F_{ca}$  为通道加权后的输出特征图。

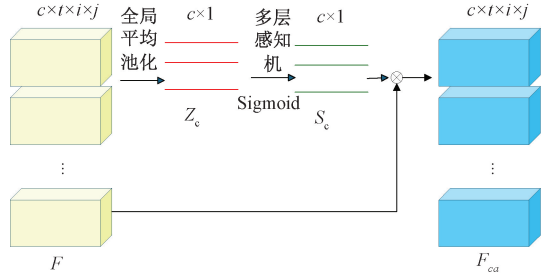


图 4 CA 模块结构

Fig. 4 CA module structure diagram

### 1.5 横向连接融合

空间分支输出特征通道数为 64, 时序分支输出特征通道数为 32; 时序分支时间维度分辨率为空间分支的 1/2, 二者在通道与时序维度上均存在差异, 无法直接融合, 因此设计横向连接融合模块, 实现双分支特征的维度对齐与互补融合。该模块仅对时序分支特征做维度变换, 不改变空间分支特征, 最终通过逐元素相加实现特征融合。其表达式如下:

$$F_{align} = \text{Conv3d}(F_{mamba}, K_{align}, S_t, P_t) \quad (10)$$

$$F_{fusion} = F_{sep} + F_{align} \quad (11)$$

式中:  $F_{mamba}$ 、 $F_{sep}$  分别为为时序分支、空间分支的输出特征图;  $K_{align}$  为  $1 \times 3 \times 3$  的对齐卷积核;  $S_t$ 、 $P_t$  分别为时间维度的卷积步长与填充大小且  $S_t = 2$ ,  $P_t = 1$ ;  $F_{align}$  为维度对齐后的时序分支特征图;  $F_{fusion}$  为双分支特征融合后的输出特征图。

### 1.6 损失函数

rPPG 任务的核心目标是使模型预测的脉搏波信号与真实信号的时序模式高度对齐, 因此本文采用负皮尔逊相关系数作为模型的损失函数。

设模型预测的 rPPG 信号为  $\hat{y} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_T]$ , 真

实 rPPG 信号为  $\mathbf{y} = [y_1, y_2, \dots, y_T]$ ,  $T$  为信号长度, 则负皮尔逊损失的数学定义为:

$$\mathcal{L} = 1 -$$

$$\frac{T \sum_{t=1}^T \hat{y}_t y_t - \left( \sum_{t=1}^T \hat{y}_t \right) \left( \sum_{t=1}^T y_t \right)}{\sqrt{T \sum_{t=1}^T \hat{y}_t^2 - \left( \sum_{t=1}^T \hat{y}_t \right)^2} \times \sqrt{T \sum_{t=1}^T y_t^2 - \left( \sum_{t=1}^T y_t \right)^2}} \quad (12)$$

## 1.7 模型结构框架

rPPG-Mamba 是一种面向非接触式生理信号测量的端到端 3D 时空特征学习模型, 其框架如图 5 所示。

首先通过帧差归一化预处理对输入 RGB 面部视频完成通道升维、局部时空特征提取与空间下采样, 输出高质量基础特征图; 随后将其分别输入空间与时序双分支进行差异化初始化, 在空间分支输入经 TDC 对时序特征增强, 然后经展平、归一化后经过深度卷积和逐

点卷积对局部细节特征提取, 在归一化、重塑特征后经 CA 模块抑制冗余特征后输出; 在时序分支输入同样先经 TDC 加强时序特征, 然后在线性化后通过时序前向和后向滑动模型提取长时序特征, 最后经线性化、CA 模块抑制冗余特征输出。空间分支聚焦局部时空特征提取, 时序分支适配长时序建模需求, 二者在三阶段递进式建模中, 前两阶段通过横向连接完成特征对齐与融合, 生成兼具局部细节与长时序依赖的互补特征, 第三阶段独立深化特征表达; 之后将双分支深层特征在通道维度拼接, 经时间上采样与通道压缩得到聚合特征, 最终通过对聚合特征进行自适应时空平均池化、 $1 \times 1 \times 1$  卷积降维与特征重塑输出 1D 的 rPPG 脉搏波信号, 训练阶段以负皮尔逊相关系数为损失函数, 引导模型参数迭代优化, 实现预测信号与真实信号在波形、周期与相位上的高度匹配。

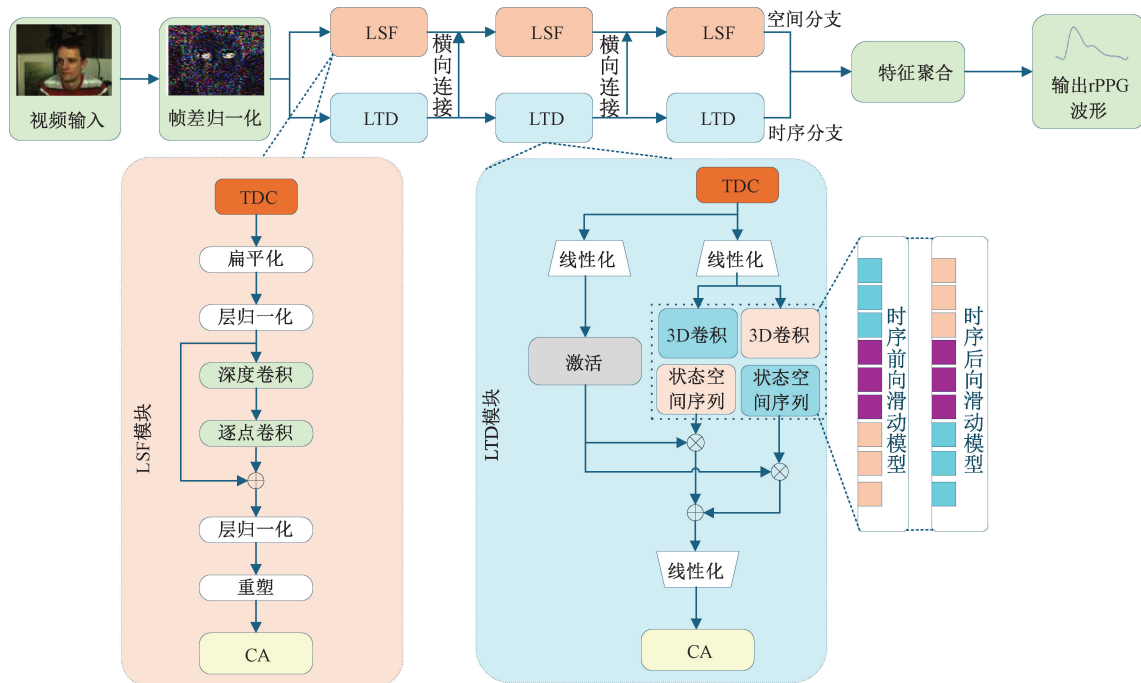


图 5 rPPG-Mamba 模型的框架

Fig. 5 Framework diagram of the rPPG-Mamba model

## 2 实验结果

### 2.1 实验设置与评价指标

为全面评估所提模型在非接触式生理参数测量任务中的性能, 本文在多个公开数据集上开展了定量实验, 包括 PURE、UBFC-rPPG 和 MMPD 数据集。采用平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE) 和皮尔逊相关系数 (pearson

correlation coefficient,  $\rho$ ) 作为评价指标, 分别衡量预测值与真实值之间的误差水平与线性相关性。对比方法包括 TS-CAN、PhysNet、DeepPhys、EfficientPhys、PhysFormer 及 PhysMamba 等主流 rPPG 模型。

### 2.2 数据集内评估结果

在 PURE 和 UBFC-rPPG 数据集上进行数据集内训练与测试, 结果如表 1 所示。

本文所提模型 (rPPG-Mamba) 在两个数据集上均取得了最优性能。在 PURE 数据集上, rPPG-Mamba 模型的

表 1 PURE 和 UBFC-rPPG 数据集内测试结果

Table 1 Intra-dataset test results of PURE and UBFC-rPPG

| 方法                            | PURE        |             |             | UBFC-rPPG   |             |             |
|-------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                               | MAE         | RMSE        | $\rho$      | MAE         | RMSE        | $\rho$      |
| TS-CAN <sup>[16]</sup>        | 2.48        | 9.01        | 0.99        | 1.70        | 2.72        | 0.99        |
| PhysNet <sup>[14]</sup>       | 2.10        | 2.60        | 0.99        | 2.95        | 3.67        | 0.97        |
| DeepPhys <sup>[13]</sup>      | 0.83        | 1.54        | 0.99        | 6.27        | 10.82       | 0.65        |
| EfficientPhys <sup>[22]</sup> | —           | —           | —           | 1.14        | 1.81        | 0.99        |
| PhysFormer <sup>[18]</sup>    | 1.10        | 1.75        | 0.99        | 0.50        | 0.71        | 0.99        |
| PhysMamba <sup>[26]</sup>     | 0.25        | 0.40        | 0.99        | 0.54        | 0.76        | 0.99        |
| rPPG-Mamba                    | <b>0.18</b> | <b>0.31</b> | <b>0.99</b> | <b>0.43</b> | <b>0.68</b> | <b>0.99</b> |

MAE 为 0.18, RMSE 为 0.31,  $\rho$  为 0.99, PhysMamba 的 MAE 为 0.25, RMSE 为 0.4, 以 MAE 为评价指标计算, 本文模型信号提取精度相对 PhysMamba 提升约 28%, 综合表现优于所有对比算法。在 UBFC-rPPG 数据集上, rPPG-Mamba 模型的 MAE 为 0.43, RMSE 为 0.68,  $\rho$  为 0.99, 略优于 PhysMamba (MAE = 0.54, RMSE = 0.76), 进一步验证了模型在常规场景下的鲁棒性。上述结果表明, rPPG-Mamba 模型在数据集内评估中具备优异的信号重建能力和心率估计精度, 具备较强的拟合能力和稳定性。

2.3 跨数据集泛化能力评估

为验证模型的泛化能力, 本文设计了跨数据集实验, 分别在 UBFC-rPPG 和 PURE 数据集上训练, 并在其他数据集上进行测试。

在 UBFC-rPPG 数据集上训练, 测试于 PURE 与 MMPD 数据集的结果如表 2 所示。

表 2 在 UBFC-rPPG 训练的跨数据集测试结果

Table 2 Cross-dataset test results of UBFC-rPPG training

| 方法                            | PURE        |             |             | MMPD         |              |             |
|-------------------------------|-------------|-------------|-------------|--------------|--------------|-------------|
|                               | MAE         | RMSE        | $\rho$      | MAE          | RMSE         | $\rho$      |
| TS-CAN <sup>[16]</sup>        | 3.69        | 13.80       | 0.82        | 14.01        | 21.04        | 0.24        |
| PhysNet <sup>[14]</sup>       | 8.06        | 19.71       | 0.61        | 12.53        | 19.27        | 0.21        |
| DeepPhys <sup>[13]</sup>      | 5.54        | 18.51       | 0.66        | 17.50        | 25.00        | 0.05        |
| EfficientPhys <sup>[22]</sup> | 5.47        | 17.04       | 0.71        | 13.78        | 22.25        | 0.09        |
| PhysFormer <sup>[18]</sup>    | 12.92       | 24.36       | 0.47        | 12.73        | 18.79        | 0.17        |
| PhysMamba <sup>[26]</sup>     | 1.20        | 5.99        | 0.97        | <b>11.96</b> | <b>17.69</b> | <b>0.29</b> |
| rPPG-Mamba                    | <b>1.18</b> | <b>5.62</b> | <b>0.97</b> | 12.46        | 18.56        | 0.27        |

在 PURE 数据集上测试时, rPPG-Mamba 模型取得了较好的测试结果, MAE 为 1.18, RMSE 为 5.62,  $\rho$  为 0.97, 优于所有的对比方法。在有极端光照与运动干扰的 MMPD 数据集上, 所有模型性能普遍下降, rPPG-Mamba 模型的 MAE 为 12.46, RMSE 为 18.56,  $\rho$  为 0.27, 大幅优于 DeepPhys 的性能 (MAE = 17.5, RMSE = 25,  $\rho$  = 0.05), 仅略逊于 PhysMamba 的性能 (MAE = 11.96, RMSE = 17.69,  $\rho$  = 0.29)。分析认为, MMPD 包含大范围

面部运动、剧烈光照突变、多肤色样本等极端干扰数据, 本文模型轻量化设计在极端噪声下特征提取鲁棒性弱于参数量更大的 PhysMamba, 因此复杂场景精度小幅落后。

在 PURE 数据集上训练, 测试于 UBFC-rPPG 与 MMPD 数据集的结果如表 3 所示。在 UBFC-rPPG 数据集上测试时, rPPG-Mamba 模型表现稳定, MAE 为 0.96, RMSE 为 1.79,  $\rho$  为 0.99, 与 PhysMamba (MAE = 0.97) 基本持平。

表 3 在 PURE 训练的跨数据集测试结果

Table 3 Cross-dataset test results of PURE training

| 方法                            | UBFC-rPPG   |             |             | MMPD         |              |             |
|-------------------------------|-------------|-------------|-------------|--------------|--------------|-------------|
|                               | MAE         | RMSE        | $\rho$      | MAE          | RMSE         | $\rho$      |
| TS-CAN <sup>[16]</sup>        | 1.30        | 2.87        | 0.99        | 13.94        | 21.61        | 0.20        |
| PhysNet <sup>[14]</sup>       | 0.98        | 2.48        | 0.99        | 13.93        | 20.29        | 0.17        |
| DeepPhys <sup>[13]</sup>      | 1.21        | 2.90        | 0.99        | 16.92        | 24.61        | 0.05        |
| EfficientPhys <sup>[22]</sup> | 2.07        | 6.32        | 0.94        | 14.03        | 21.62        | 0.17        |
| PhysFormer <sup>[18]</sup>    | 1.44        | 3.77        | 0.98        | 14.57        | 20.71        | 0.15        |
| PhysMamba <sup>[26]</sup>     | 0.97        | 1.93        | 0.99        | <b>10.31</b> | <b>16.02</b> | <b>0.34</b> |
| rPPG-Mamba                    | <b>0.96</b> | <b>1.79</b> | <b>0.99</b> | 11.80        | 17.85        | 0.26        |

在 MMPD 数据集上测试, rPPG-Mamba 模型的 MAE 为 11.8, RMSE 为 17.85,  $\rho$  为 0.26, 与 DeepPhys (MAE = 16.92, RMSE = 24.61) 相比性能大幅提升, 仅略逊于 PhysMamba 的性能 (MAE = 10.31, RMSE = 16.02,  $\rho$  = 0.34)。轻量化结构参数容量有限, 面对训练集未见的极端扰动泛化性能不及大参数量 PhysMamba, 但相较 DeepPhys 等传统卷积模型仍具备显著优势。综上, rPPG-Mamba 模型在跨数据集任务中表现出较强的迁移能力, 但在复杂干扰场景下仍有优化空间。

2.4 消融实验分析

为验证各模块的有效性, 本文在 UBFC-rPPG 数据集上开展了消融实验, 结果如表 4 所示。

表 4 主要模块的消融实验

Table 4 Ablation experiments of major modules

| 方法                   | UBFC-rPPG   |             |             |
|----------------------|-------------|-------------|-------------|
|                      | MAE         | RMSE        | $\rho$      |
| 移除 TDC               | 0.64        | 0.83        | 0.84        |
| 移除 CA                | 0.61        | 0.78        | 0.87        |
| 移除 Bi-Mamba          | 0.57        | 0.76        | 0.91        |
| Bi-Mamba 替换为单向 Mamba | 0.55        | 0.74        | 0.93        |
| 仅空间分支                | 0.53        | 0.79        | 0.81        |
| 仅时序分支                | 0.51        | 0.72        | 0.83        |
| 移除横向连接               | 0.48        | 0.94        | 0.92        |
| rPPG-Mamba           | <b>0.43</b> | <b>0.68</b> | <b>0.99</b> |

移除 TDC 模块: MAE 上升至 0.64, RMSE 为 0.83, 皮尔逊相关系数  $\rho$  降至 0.84, 性能下降显著, 说明 TDC 模块对信号细节建模具有重要作用, 能够有效捕捉局部时序特征并抑制噪声干扰。

移除 CA 模块: MAE 为 0.61, RMSE 为 0.78,  $\rho$  为 0.87, 性能下降幅度较大, 表明注意力机制有助于增强特征表达能力。

移除双向 Mamba 模块: MAE 上升至 0.57, RMSE 为 0.76,  $\rho$  为 0.91, 验证了 Bi-Mamba 模块在长时序依赖建模中的核心地位, 双向结构能够充分挖掘 rPPG 信号的前后上下文关联, 能有效提升精度。

为进一步验证双向 Mamba 的必要性, 设计了单向 Mamba 替换双向 Mamba, MAE 为 0.55, RMSE 为 0.74,  $\rho$  为 0.93。由消融结果可知, 单向 Mamba 架构受时序信息丢失影响, 检测精度弱于双向 Mamba, 但优于去除 Mamba 模块的性能。充分验证了双向 Mamba 的必要性, 双向时序建模能够稳定提升微弱信号的提取精度。

仅使用空间分支与仅使用时序分支, 各项指标均劣于完整模型, 说明仅使用深度可分离卷积或仅使用双向 Mamba 的单分支模型均存在明显性能短板, 表明空间分支擅长局部时空特征提取, 时序分支专注长时序依赖建模, 两大分支功能互补, 不存在功能重叠问题, 两者协同作用效果更优。

移除横向连接, MAE 仅上升至 0.48, 但 RMSE 大幅升至 0.9,  $\rho$  为 0.92, 说明多尺度特征融合对提升预测稳定性具有积极影响。

完整模型 rPPG-Mamba 在所有指标上均取得最优表现: MAE=0.43, RMSE=0.68,  $\rho=0.99$ , 充分证明了各模块协同作用的有效性, 实现了精度、稳定性与信号拟合度的全面提升。

2.5 参数与计算效率

为全面评估所提 rPPG-Mamba 模型的设备端部署实用性, 本文从参数量与乘法累加运算量两个维度, 与当前主流 rPPG 算法进行了量化对比, 结果如表 5 所示。表 5 中所有对比模型统一采用输入视频尺寸 128×128 pixels、单段视频帧长 160 帧, 保证参数量与运算量对比条件

一致。

表 5 参数量和运算量的比较结果

Table 5 Comparison results of parameters and computational complexity

| 方法                            | 参数量/( $\times 10^6$ ) | 乘法累加运算量/GMac |
|-------------------------------|-----------------------|--------------|
| TS-CAN <sup>[16]</sup>        | 7.50                  | 96.0         |
| PhysNet <sup>[14]</sup>       | 0.77                  | 56.1         |
| DeepPhys <sup>[13]</sup>      | 7.50                  | 96.0         |
| EfficientPhys <sup>[22]</sup> | 7.40                  | 45.6         |
| PhysFormer <sup>[18]</sup>    | 7.38                  | 40.5         |
| PhysMamba <sup>[26]</sup>     | 0.56                  | 47.3         |
| rPPG-Mamba                    | <b>0.45</b>           | <b>37.4</b>  |

实验结果表明, 本文提出的 rPPG-Mamba 模型不仅保持极低计算开销, 还大量减少了参数规模。与 TS-CAN 和 DeepPhys 相比, rPPG-Mamba 的参数量仅为其约 10%; 与 EfficientPhys 相比, 参数量降低了约 89.9%。这一显著优势主要得益于模型的双分支轻量化设计, 在保留核心时空特征建模能力的前提下, 有效规避了传统深度学习模型参数冗余的问题。在计算复杂度方面, rPPG-Mamba 远低于 PhysNet 与 TS-CAN 等方法。这表示本模型在保证推理速度的同时, 极大地降低了硬件资源消耗。综上所述, rPPG-Mamba 模型通过双分支轻量化架构, 实现了高精度与低开销的平衡, 具备优异的设备部署与实时监测应用潜力。

2.6 信号重建效果与心率预测精度

模型的信号拟合能力如图 6 所示, 其中预测信号与真实信号的波形趋势、峰谷位置高度一致, 相位匹配度良好。结果表明, 所提模型能够有效重建出与真实 PPG 信号高度吻合的波形, 具备较强的去噪能力与时序特征捕捉能力, 为非接触式远程生理测量提供了可靠的信号基础。

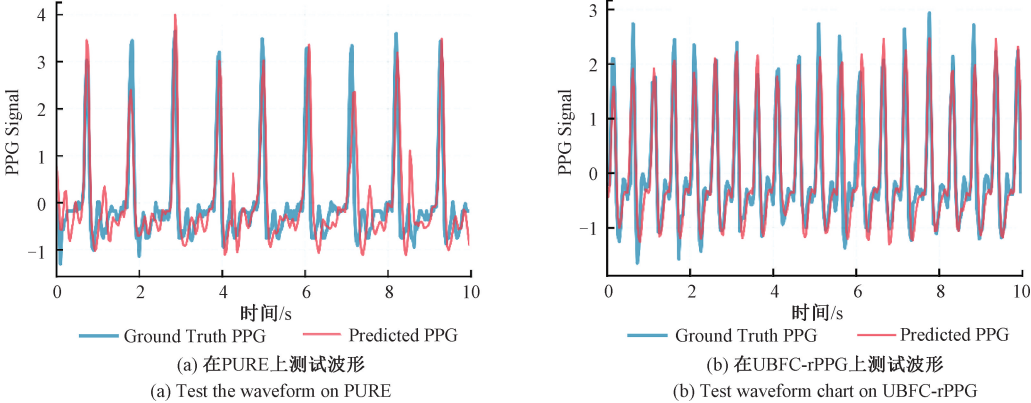


图 6 在 PURE 和 UBFC-rPPG 上跨数据集测试的 rPPG 预测信号与真实值示例曲线

Fig. 6 Example curves of rPPG predicted signals versus ground truth in cross-dataset testing on PURE and UBFC-rPPG

模型预测心率与真实心率相关性如图 7 所示,其中散点非常紧密地分布在对角参考线的两侧,在 PURE 和 UBFC-rPPG 数据集上  $\rho$  分别为 0.97 和 0.99,说明预测

值和真实值之间具有极强的线性相关性。结果表明,模型在两个数据集上均实现了高相关性心率预测,具备良好的泛化性能与实际应用潜力。

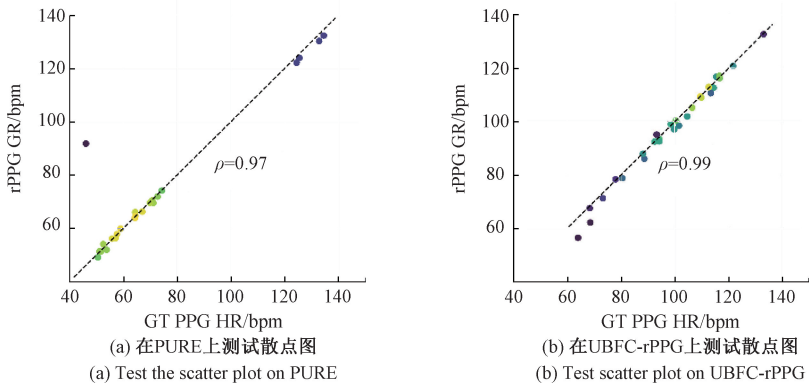


图 7 在 PURE 和 UBFC-rPPG 上跨数据集测试的散点图

Fig. 7 Scatter plot of cross-dataset testing on PURE and UBFC-rPPG

3 结 论

针对现有 rPPG 提取算法存在的测量精度与轻量化部署难以兼顾、长时序依赖建模效率不足等问题,本文提出了一种融合深度可分离卷积和双向 Mamba 的端到端 rPPG-Mamba 模型,该模型采用双分支架构,通过空间分支提取局部时空特征,时序分支建模长距离时空依赖,并结合横向连接实现多尺度特征融合,有效平衡了测量精度与计算效率。在 PURE、UBFC-rPPG 和 MMPD 数据集上的实验结果表明,数据集内评估中,所提模型在数据集内评估中均取得最优性能;跨数据集实验中,模型在常规场景下表现稳定,性能较对比模型有显著提升,仅在极端干扰场景下的 MMPD 数据集略低于 PhysMamba;且参数量与计算量显著降低,提升了轻量化部署可行性,为端侧设备的实时非接触式生理测量提供了支撑。本文仍存在一些不足,模型在运动幅度大、暗光环境下泛化性能受限,极端场景精度出现明显下滑;当前模型仅实现单一心率参数提取,无法同步估算血氧、呼吸等多类生理指标。

综上,rPPG-Mamba 模型在保证高测量精度的同时具备良好的泛化能力与轻量化部署潜力,为非接触式生理测量技术的实际应用提供了有效解决方案。未来工作将聚焦于复杂场景鲁棒性增强、多参数同步监测。

参考文献

[ 1 ] Wei Lan, Tian Yonghong, Wang Yaowei, et al. Automatic webcam-based human heart rate measurements using Laplacian eigenmap [ C ] // Proceedings of the 11th Asian Conference on Computer Vision. Berlin: Springer, 2012: 281-292.

[ 2 ] 淳新益, 郑秀娟, 张畅, 等. 遗传算法优化 BP 神经网络的非接触式血压估计方法 [ J ]. 电子测量与仪器学报, 2021, 35(7): 54-59.  
Chun Xinyi, Zheng Xiujuan, Zhang Chang, et al. Non-contact blood pressure estimation method based on genetic algorithm optimized BP neural network [ J ]. Journal of Electronic Measurement and Instrumentation, 2021, 35(7): 54-59.

[ 3 ] Niu Xuesong, Han Hu, Shan Shiguang, et al. VIPL-HR: A multimodal database for pulse estimation from less-constrained face video [ C ] // Proceedings of the 14th Asian Conference on Computer Vision. Berlin: Springer, 2018: 611-627.

[ 4 ] Van Gastel M, Stuijk S, De Haan G. Motion robust remote PPG in infrared [ J ]. IEEE Transactions on Biomedical Engineering, 2015, 62(5): 1425-1433.

[ 5 ] Tsouri G R, Kyal S, Dianat S A, et al. Constrained independent component analysis approach to nonobtrusive pulse rate measurements [ J ]. Journal of Biomedical Optics, 2012, 17(7): 077011.

[ 6 ] Verkruyse W, Svaasand L O, Nelson J S. Remote plethysmographic imaging using ambient light [ J ]. Optics Express, 2008, 16(26): 21434-21445.

[ 7 ] 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.

[ 8 ] De Haan G, Jeanne V. Robust pulse rate from chrominance-based rPPG [ J ]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2878-2886.

[ 9 ] Wang W, Stuijk S, De Haan G. A novel algorithm for

- remote photoplethysmography: Spatial subspace rotation[J]. *IEEE Transactions on Biomedical Engineering*, 2015, 63(9): 1974-1985.
- [10] Liu Mingxuan, Tang Jiankai, Chen Yongli, et al. Spiking PhysFormer: Camera-based remote photoplethysmography with parallel spike driven transformer [PP/OL]. V4. (2025-01-03) [2026-04-14]. <https://doi.org/10.48550/arXiv.2402.04798>.
- [11] 朱凌建, 陈剑虹, 王裕鑫, 等. 基于 GRU 神经网络的脉搏波波形预测方法研究[J]. *电子测量与仪器学报*, 2022, 36(5): 242-248.
- Zhu Lingjian, Chen Jianhong, Wang Yuxin, et al. Research on pulse wave waveform prediction method based on GRU neural network[J]. *Journal of Electronic Measurement and Instrumentation*, 2022, 36(5): 242-248.
- [12] Li Kunchang, Li Xinhao, Wang Yi, et al. VideoMamba: State space model for efficient video understanding [PP/OL]. V2. (2024-03-12) [2026-04-14]. <https://doi.org/10.48550/arXiv.2403.06977>.
- [13] 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.
- [14] Yu Zitong, Li Xiaobai, Zhao Guoying, et al. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks [C]//*Proceedings of the British Machine Vision Conference*. Cardiff: BMVA Press, 2019: 277.
- [15] Niu Xuesong, Shan Shiguang, Han Hu, et al. RhythmNet: End-to-end heart rate estimation from face video via spatial-temporal representation learning [J]. *IEEE Internet of Things Journal*, 2020, 7(12): 11341-11351.
- [16] Paruchuri A, Liu Xin, Pan Yulu, et al. Motion matters: Neural motion transfer for better camera physiological measurement [PP/OL]. V4. (2023-11-06) [2026-04-14]. <https://doi.org/10.48550/arXiv.2303.12059>.
- [17] Yu Sungnien, Wang Chienshun, Chang Yuping. Heart rate estimation from remote photoplethysmography based on lightweight U-Net and attention modules [J]. *IEEE Access*, 2023, 11: 67892-67901.
- [18] Yu Zitong, Shen Yuming, Shi Jingang, et al. PhysFormer: Facial video-based physiological measurement with temporal difference transformer [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, 2022: 1164-1173.
- [19] Bousefsaf F, Pruski A, Maaoui C. 3D convolutional neural networks for remote pulse rate measurement and mapping from facial video [J]. *Applied Sciences*, 2019, 9(20): 4364.
- [20] Tulyakov S, Alameda-Pineda X, Ricci E, et al. Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2016: 2396-2404.
- [21] Ouzar Y, Djeldji D, Bousefsaf F, et al. X-iPPGNet: A novel one-stage deep learning architecture based on depthwise separable convolutions for video-based pulse rate estimation [J]. *Computers in Biology and Medicine*, 2023, 154: 106592.
- [22] Liu Xin, Hill B, Jiang Ziheng, et al. EfficientPhys: Enabling simple, fast and accurate camera-based cardiac measurement [C]//*Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. IEEE, 2023: 4155-4164.
- [23] 孙苏云, 郭剑, 付阳烨, 等. 基于 FMCW 雷达的非接触式心率估计方法 [J]. *电子测量技术*, 2024, 46(14): 117.
- Sun Suyun, Guo Jian, Fu Yangye, et al. Non-contact heart rate estimation method based on FMCW radar [J]. *Electronic Measurement Technology*, 2024, 46(14): 117.
- [24] Ma Jun, Li Feifei, Wang Bo. U-Mamba: Enhancing long-range dependency for biomedical image segmentation [PP/OL]. V1. (2024-01-09) [2026-04-14]. <https://doi.org/10.48550/arXiv.2401.04722>.
- [25] Gu A, Dao T. Mamba: Linear-time sequence modeling with selective state spaces [PP/OL]. V2. (2024-05-31) [2026-04-14]. <https://doi.org/10.48550/arXiv.2312.00752>.
- [26] Liu X, Paruchuri A, Narayanswamy G, et al. PhysMamba: Synergistic state space duality model for remote physiological measurement [C]//*International Conference on Artificial Neural Networks*. Berlin: Springer, 2025: 16071:128-140.
- [27] Qian Jian, Goh T L, Xie Bingyu, et al. BioMamba: Leveraging spectro-temporal embedding in bidirectional mamba for enhanced biosignal classification [PP/OL]. V3. (2025-04-25) [2026-04-14]. <https://doi.org/10.48550/arXiv.2503.11741>.
- [28] Wang Xian, Zeng Wenzhi, Xu Junzeng, et al. MM-3D UNet: Development of a lightweight breast cancer tumor segmentation network utilizing multi-task and depthwise separable convolution [J]. *Frontiers in Oncology*, 2025, 15: 1563959.
- [29] 王宇, 戎舟, 王苏煜. 基于 3D 残差注意力网络的非接触式心率测量方法 [J]. *国外电子测量技术*, 2024,

42(2): 53-59.

Wang Yu, Rong Zhou, Wang Suyu. Non-contact heart rate measurement method based on 3D residual attention network [J]. Foreign Electronic Measurement Technology, 2024, 42(2): 53-59.

[30] Bobbia S, Macwan R, Benezeth Y, et al. Unsupervised skin tissue segmentation for remote photoplethysmography [J]. Pattern Recognition Letters, 2019, 126: 107-114.

[31] 成娟, 武梦影, 宋仁成, 等. 模型和盲源分离融合的运动鲁棒视频心率检测[J]. 仪器仪表学报, 2023, 44(8): 289-298.

Cheng Jun, Wu Mengying, Song Rencheng, et al. Motion robust video heart rate detection based on model and blind source separation fusion[J]. Chinese Journal of Scientific Instrument, 2023, 44(8): 289-298.

[32] Chollet F. Xception: Deep learning with depthwise separable convolutions [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 1251-1258.

[33] Liu X, Narayanswamy G, Paruchuri A, et al. rPPG-Toolbox: Deep remote ppg toolbox [C]//Advances in Neural Information Processing Systems. NeurIPS, 2023: 28763-28775.

[34] 童凯南, 林友芳, 刘军, 等. 面向时空交通栅格流量预测的 3D 通道注意力网络[J]. 国防科技大学学报, 2022, 44(3): 1-8.

Tong Kainan, Lin Youfang, Liu Jun, et al. 3D channel attention network for spatiotemporal traffic grid flow prediction[J]. Journal of National University of Defense Technology, 2022, 44(3): 1-8.

作者简介



**朱海军**, 2021 年于河北工业大学获得博士学位, 现为河北大学电子信息工程学院讲师, 主要研究方向为神经调控与生理信号分析。

E-mail: suseyinxiang@126.com

**Zhu Haijun** received his Ph. D. degree from Hebei University of Technology in 2021. Now he is a lecturer in the School of Electronic Information Engineering at Hebei University. His main research interests include neural regulation and physiological signal analysis.



**杜硕**(通信作者), 2024 年于东北大学获得博士学位, 现为河北大学电子信息工程学院讲师, 主要研究方向为心血管系统建模与非接触式生理测量。

E-mail: dushuo@hbu.edu.cn

**Du Shuo** (Corresponding author) received her Ph. D. degree from Northeastern University in 2024. Now she is a lecturer in the School of Electronic Information Engineering at Hebei University. Her main research interests include cardiovascular system modeling and non-contact physiological measurement.