

基于可分离卷积的人脸视频心率检测方法

王 猛, 杨观赐

(贵州大学 现代制造技术教育部重点实验室, 贵州 贵阳 550025)

摘 要: 心率估计在分析人体健康状态方面具有重要作用。针对基于远程非接触式心率检测方法存在的脸部特征建模不充分问题,提出了一种基于深度可分离卷积的人脸视频心率检测方法(FVSC-HR)。首先,通过融入卷积门控单元,获取相邻帧的细粒度特征;其次,采用基于时间差分卷积的多头注意力机制,细化局部时空表示的抗干扰能力;最后,通过时空前馈网络输出心率值。在UBFC-rPPG和UBFC-phys数据集上的实验结果表明,FVSC-HR方法在UBFC-rPPG数据集上的MAE为3.4, ρ 评价指标为0.8;在UBFC-phys数据集上的MAE为7.3。与现有的6种方法相比,FVSC-HR的性能显著更优,可为智能健康监测、远程医疗等领域提供高效的心率检测技术支持。

关键词: 心率检测;人脸视频;可分离卷积;差分注意力机制;卷积门控单元

DOI:10.11907/rjdk.241526

开放科学(资源服务)标识码(OSID):



中图分类号:TP18

文献标识码:A

文章编号:1672-7800(2025)008-0038-05

Separable Convolution-based Heart Rate Detection Method for Face Video

WANG Meng, YANG Guanci

(Key Laboratory of Advanced Manufacturing Technology of the Ministry of Education, Guizhou University, Guiyang 550025, China)

Abstract: Heart rate estimation plays an important role in analyzing human health status such as heart. In order to address the problem of inadequate face feature modeling in remote contactless heart rate detection, the separable convolution based heart rate detection for face video (FVSC-HR) is proposed. First, the convolutional gating unit is incorporated to obtain fine-grained features of neighboring frames. Second, a multi-head attention mechanism based on temporal differential convolution is used to refine the local spatio-temporal representation anti-interference capability. Then, the heart rate is output through the spatio-temporal feed-forward network. The model is trained and tested on UBFC-rPPG and UBFC-phys datasets. Tested on the UBFC-rPPG and UBFC-phys datasets, the MAE of the FVSC-HR method on the UBFC-rPPG dataset is 3.4, and the ρ -evaluation index is 0.8; on the UBFC-phys dataset, the MAE of the FVSC-HR method is 7.3; the FVSC-HR outperforms the six methods compared, and it can be used to provide intelligent health monitoring and telemedicine to provide heart rate detection technology.

Key Words: heart rate detection; face video; separable convolution; differential attention mechanism; convolutional gating unit

0 引言

心率等生理信号通常可以评估一个人的健康状况,预测潜在疾病的风险。心率测量可由心率测量仪进行测量,这种方法虽然测量精准,但接触式测量设备存在一些缺点,如长期使用可能导致不适,且皮肤温度、颜色和患者运

动等因素也会影响其测量的准确性^[1]。因此,研究人员开始探索非接触式的远程心率检测方法。目前,远程光电容积脉搏波方法(Remote photoplethysmography, rPPG)不需要佩戴接触装置,不但避免了接触装置的各种弊端,而且还适用于长期连续监测,但是光照条件、皮肤等各种因素都会影响rPPG方法的准确性。因此,需要一种细化脸部特征和全局聚合特征信息的远程心率检测方法。

收稿日期:2024-06-13

扫描二维码阅读全文:



基金项目:国家自然科学基金项目(62163007,62373116);贵州省科技计划项目(黔科合平台人才[2020]6007-2,黔科合支撑[2023]一般118)

作者简介:王猛(1998-),男,贵州大学现代制造技术教育部重点实验室硕士研究生,研究方向为自主智能系统与机器人;杨观赐(1983-),男,博士,CCF会员,贵州大学现代制造技术教育部重点实验室教授、博士生导师,研究方向为自主智能系统与机器人、多模态数据融合认知计算。本文通讯作者:杨观赐。

远程光电容积脉搏波方法主要分为常规rPPG方法和基于深度学习的rPPG方法。在深度学习盛行之前,常规rPPG方法主要是消除面部视频中产生的运动伪影和噪声。文献[2]提出了环境光和数码相机远程测量脉搏信号的方法,并通过实验表明该方法可以估计心脏和呼吸速率,帮助医生进行辅助诊断。但不足的是此方法受限于采集视频的传感器。文献[3]提出了基于色度的方法,提高了模型的鲁棒性,遗憾的是所提出的方法测试样本较少,难以反映其泛化性。为此,Wang等^[4]提出了结合皮肤反射和生理学特性的数学模型,设计特定应用的rPPG解决方案,提高模型的泛化性。文献[5]研究先验知识对面部视频估计任务的影响,引入可微局部变换的特征不变量,提高了模型在室外视频的心率估计性能。但其不足的是提出的心率检测方法预测心率的总体误差较大,难以在实际场景中进行广泛应用。

随着深度学习的发展,基于深度学习的远程心率估计方法成为主流。为提高rPPG估计心率的准确率,黄海等^[6]提出基于多头自注意力机制的心率估计模型,设计时空表示转换方法,抑制非感兴趣区域引入的噪声。王宇等^[7]提出基于3D残差注意力网络的心率估计方法,融合3D卷积注意力机制和时空卷积,有利于模型提取通道和空间特征信息,关注生理信号丰富的区域。但不足的是该方法未考虑复杂条件下光照等因素对心率检测所产生的影响。Yu等^[8]提出自适应聚合局部和全局时空特征的端到端transformer架构,设计时间差分注意力机制,增强周期rPPG特征。Yu等^[9]提出基于双通道的时间差分注意力机制,融合时间差分周期和交叉注意力变换,更好地利用时间上下文信息和周期rPPG线索。为降低光照带来的影响,Wang等^[10]构建心率鲁棒测量的自监督机器学习网络,在不使用任何标签的情况下进行训练并取得理想的效果。但不足的是模型参数较为复杂,计算量大。为解决受限于大量数据和抗噪声能力不足的问题,Liu等^[11]提出了自相似先验的线性自监督重建模型,设计噪声不敏感策略,减少运动和光照的干扰。

总体而言,基于远程视频的心率检测方法已展现出显著的性能优势,但其性能仍受到光照条件、肤色差异以及视频中脸部角度变化等因素的限制。此外,如何捕捉脸部微弱特征信息、聚合特征以及将特征转化为PPG信号信息,已成为远程光电容积脉搏波描记法的关键。

本文受TransNeXt模型^[12]中卷积门控线性单元能够提高模型全局信息聚合质量和鲁棒性的启发,提出了一种基于可分离卷积的人脸视频心率检测方法(Seperable Convolution Based Heart Rate Detection for Face Video, FVSC-HR)。

1 基于可分离卷积的人脸视频心率检测方法

心率检测是分析人体健康状况的重要途径,而日常心

率数据的监测尤为重要。为通过心率检测分析健康状态,本文提出了基于可分离卷积的人脸视频心率检测方法。首先,融入卷积门控线性单元,以获取相邻帧的细粒度特征;其次,采用基于时间差分卷积的多头注意力机制,增强周期脸部特征信息,细化局部时空表示抗干扰能力;最后,通过时空前馈网络输出心率。FVSC-HR模型主要由卷积门控线性单元、时间差分多头注意力机制和时空前馈网络3个部分组成。FVSC-HR模型的架构如图1所示。

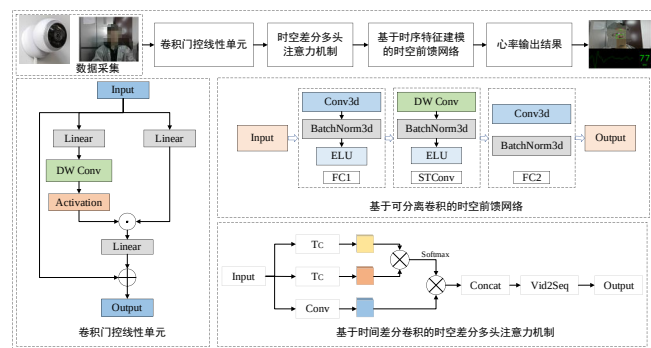


Fig. 1 Architecture of separable convolution based heart rate detection for face video

图1 基于可分离卷积的人脸视频心率检测方法架构

1.1 基于最近邻特征的卷积门控线性单元

门控线性单元(Gated Linear Unit, GLU)由2个线性投影组成。然而,由于GLU在特征图上的所有令牌共享相同的门控信号,导致其信道注意力缺乏灵活性。为增强局部建模能力,文献[12]提出卷积门控线性单元,通过将可分离卷积引入GLU的门控分支激活函数中,使模型能够捕获最近邻图像特征,从而形成基于最近邻特征的卷积门控线性单元。计算公式如式(1)一式(3)所示。

$$t_1 = L(f) \quad (1)$$

$$t = t_1 \cdot R(D_c(L(f))) \quad (2)$$

$$x = f + t \quad (3)$$

式中: $f(f_1, \dots, f_n)$ 是卷积门控线性单元的输入; t, t_1 表示运算过程的中间量; L 代表线性操作; R 表示激活函数; D_c 代表卷积核 3×3 的深度可分离卷积; $x(x_1, \dots, x_n)$ 表示卷积门控线性单元的输出矩阵; n 表示视频帧的数量。

1.2 基于时间差分卷积的时空差分多头注意力机制

目前,面部心率检测方法侧重于使用有限时空感受野的卷积神经网络来挖掘微妙的PPG线索,这忽略了面部建模的长时空感知和交互^[8]。因此,本文采用基于时间差分卷积的时空差分多头注意力机制,引导全局注意力细化局部时空特征。时空差分多头注意力机制的公式如式(4)一式(8)所示。

假设时空差分多头注意力机制的输入矩阵为 $x(x_1, \dots, x_n)$,经过时间差分卷积映射成 Q, K, V 矩阵,查询 Q 、键值 K 进行矩阵相乘,得到中间注意力矩阵 S 。之后,短矩阵 S 与矩阵 V 进行数学相乘交叉运算,得到输出矩阵 M 。

$$Q = B(T_c(x)) \quad (4)$$

$$K = B(T_c(x)) \quad (5)$$

$$V = T_c(x) \quad (6)$$

$$S = d(\text{Softmax}(Q \otimes K)) \quad (7)$$

$$M = S \otimes V \quad (8)$$

式中: Q 、 K 、 V 分别表示线性映射后的 Q 矩阵、 K 矩阵和 V 矩阵; B 表示批归一化操作; T_c 表示时间差分卷积; Softmax 表示最大批量归一化操作; d 表示随机丢包操作; S 表示 Q 和 K 相乘后的中间注意矩阵; M 表示时空差分多头注意力机制输出矩阵。

1.3 基于可分离卷积的时空前馈网络

前馈网络由2个线性变换层组成,其中隐藏层的维度被扩展,以学习更丰富的特征表示。文献[8]提出基于可分离卷积的时空前馈网络,细化局部一致性和抑制部分噪声特征,从而获取相对位置线索。基于可分离卷积的时空前馈网络架构如图2所示。

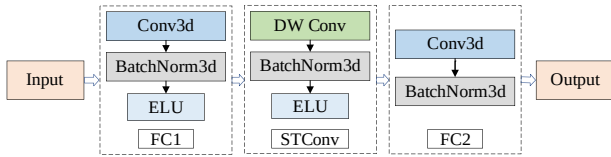


Fig. 2 Architecture of spatio-temporal feedforward network based on separable convolution

图2 基于可分离卷积的时空前馈网络架构

其中, Conv3d 代表卷积核为1的三维卷积, ELU 代表指数线性单元(Exponential Linear Unit), BatchNorm3d 代表三维批量归一化, DW Conv 代表卷积核为3的深度可分离卷积。

假设时空前馈的输入为时空差分多头注意力机制输出的 M 矩阵。首先, M 矩阵经过卷积核为1的三维卷积、批量归一化及激活函数后,得到 F_1 矩阵;其次,经过卷积核为1的深度可分离卷积、批量归一化及激活函数后,得到矩阵 S_T ,丰富携带的特征信息;最后,经过卷积核为3的三维卷积、批量归一化后,得到基于时序特征建模的时空前馈网络的输出矩阵 Z 。计算公式如式(9)一式(11)所示。

$$F_1 = \text{ELU}(B(C(M))) \quad (9)$$

$$S_T = \text{ELU}(B(D_c(F_1))) \quad (10)$$

$$Z = B(C_3(S_T)) \quad (11)$$

式中: C 代表卷积核为1的三维卷积; ELU 代表指数线性单元; B 代表三维批量归一化; D_c 代表卷积核为3的深度可分离卷积; C_3 表示卷积核为1的三维卷积; F_1 、 S_T 表示中间输出矩阵; Z 表示输出矩阵。

1.4 损失函数

FVSC-HR模型采用回归损失函数和分类损失函数来计算损失。

MSE损失函数是计算预测值与目标值之间差值平方和的均值的损失函数。计算公式如式(12)所示。

$$\text{loss}_1 = \frac{1}{n} \sum_{i=1}^n (y_i - y)^2 \quad (12)$$

式中: y_i 表示样本的预测值; y 表示当前样本的目标值; n 表示样本数量。

平均绝对值误差损失函数(Mean Absolute Error, MAE)是计算预测值与目标值之间的平均误差。计算公式如式(13)所示。

$$\text{loss}_2 = \sum_{i=1}^n |y_i - y| \quad (13)$$

交叉熵损失函数(Cross-Entropy loss)是计算预测值和真实标签值的距离度量,其计算公式如式(14)所示。

$$\text{loss}_3 = \frac{1}{n} \sum_i -(y \cdot \log(p_i) + (1 - y) \log(1 - p_i)) \quad (14)$$

式中: p_i 表示模型预测的概率; y 表示当前样本的目标值。

联合公式(12)、(13)和(14)可知,FVSC-HR模型的损失函数计算公式如式(15)所示。

$$\text{loss} = \text{loss}_1 + \text{loss}_2 + \text{loss}_3 \quad (15)$$

2 实验结果与分析

2.1 数据集与评价指标

2.1.1 数据集

采用UBFC-rPPG^[13]和UBFC-phys^[14]两个公共数据集测试模型的性能。

UBFC-rPPG数据集包含50个视频。视频使用罗技C920 HD Pro网络摄像头录制的,分辨率为640×480、帧率为30 fps。同时,通过CMS50E透射脉搏血氧仪收集实况PPG数据和心率值。UBFC-rPPG数据集包含2个子数据集,本文采用的是子数据集2。子数据集2共包含42个视频,记录心率等真实数据,便于研究人员使用。

UBFC-phys数据集包含56名参与者,这些参与者按照著名的特里尔社会压力测试启发的严格实验协议进行测试。实验过程分为休息阶段、演讲阶段以及不同难度的任务阶段三个部分。每个视频的分辨率为1 024×1 024、帧率为35fps。

2.1.2 评价指标

采用心率检测中的平均绝对误差(MAE)、均方根误差(RMSE)、Pearson相关系数(ρ)以及平均绝对百分比误差(MAPE)作为评价指标,对所提出模型的性能进行评估。

MAE是预测值与真实值之间的差距,取值范围为(0, +∞)。当预测值与真实值完全吻合时等于0,认为预测效果最好。计算公式如式(16)所示。

$$\text{MAE} = \frac{1}{M} \sum_{i=1}^M |x_i - y_i| \quad (16)$$

式中: x_i 表示第 i 个预测值; y_i 表示第 i 个真实值; M 表示样本数量。

RMSE是均方误差的平方根,其取值范围为(0, +∞)。

当预测值与真实值完全吻合时等于0,认为预测效果最好。计算公式如式(17)所示。

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (x_i - y_i)^2} \quad (17)$$

ρ 是用于两个变量之间是否线性相关。计算公式如式(18)所示。

$$\rho = \frac{\sum_{i=1}^M (x_i - x_j)(y_i - y_j)}{\sqrt{\sum_{i=1}^M (x_i - x_j)^2} \sqrt{\sum_{i=1}^M (y_i - y_j)^2}} \quad (18)$$

式中: x_j 表示样本预测值中平均值; y_j 表示样本真实值中平均值。

MAPE的取值范围为(0, +∞),当MAPE=0时,表示模型效果最好;当MAPE>100%时,表示为劣质模型。其计算公式如式(19)所示。

$$MAPE = \frac{100\%}{M} \sum_{i=1}^M \left| \frac{x_i - y_i}{y_i} \right| \quad (19)$$

2.1.3 实验参数设置

本文模型的实验均基于PyTorch框架实现,并在显存为32 GB的NVIDIA GTX v100GPU上进行训练和测试。rPPG-Toolbox^[15]是一个基于开源平台的工具箱,集成了PhysFormer^[8]、DeepPhys^[20]及PhysNet^[21]等心率检测算法。本文在rPPG-Toolbox平台上实现了FVSC-HR算法,并采用同步批处理归一化方法对模型进行训练。训练过程中,使用初始学习率为1e-4的AdamW优化器,动量设置为0.9,同时采用步长为15的固定步长衰减策略来调整学习率。本文选择MAE作为网络训练的损失函数,训练轮次设置为50,批次大小设置为8,并使用MSE作为回归损失函数。

2.2 实验结果及分析

2.2.1 单个数据集内的模型性能统计结果

为合理评估所提出模型的有效性,本文进行了实验验证。首先,对模型参数进行配置;其次,其他对比模型按照文中的参数设置进行训练和测试,实验均在UBFC-rPPG数据集上完成。同时,本文提出的FVSC-HR模型也在UBFC-rPPG数据集上进行了训练和测试,最终得到表1所示UBFC-rPPG数据集上不同模型的性能统计结果。

根据表1中的实验数据,从MAE的评价指标分析,

Table 1 Performance statistics of different models on the UBFC-rPPG dataset

方法	年份	MAE	RMSE	MAPE	ρ
POS ^[4]	2016	4.00	7.60	3.90	0.9
LGI ^[5]	2018	15.80	28.50	14.60	0.4
STSC ^[16]	2023	2.15	3.82	3.15	1.0
EEMD+LA-SSA ^[17]	2022	1.37	3.61	-	1.0
STI ^[18]	2022	3.90	6.20	-	-
DSTA-HR ^[19]	2023	4.90	7.30	-	-
FVSC-HR(Ours)	2024	3.40	10.80	3.80	0.8

“-”代表着原始论文中无相关评价指标的数据

POS模型的MAE值为4.0,LGI模型的MAE值为15.8,STSC模型的MAE值为2.15,EEMD+LA-SSA模型的MAE值为1.37,是MAE评价指标中模型取得的最佳值。本文提出的FVSC-HR模型的MAE值为3.4,在所对比的模型中位列第三。从RMSE评价指标分析,POS模型的RMSE值为7.6,LGI模型的RMSE值为28.5,STSC模型的RMSE值为3.82,EEMD+LA-SSA模型的RMSE值为3.61,是RMSE评价指标中模型取得的最佳值。FVSC-HR模型的RMSE值为10.8,超过了LGI模型。从MAPE评价指标分析,POS模型的MAPE值为3.9,LGI模型的MAPE值为14.6,STSC模型的MAPE值为3.15,是MAPE评价指标中模型取得的最佳值。FVSC-HR模型的MAPE值为3.8,在所对比的模型中位列第二。

综上,本文提出的FVSC-HR模型,其MAE评价指标为3.4, RMSE评价指标为10.8, MAPE评价指标为3.8, ρ 评价指标为0.8,在所有对比模型中展现了具有竞争力的性能。

2.2.2 跨不同数据集的模型性能统计结果

为合理评估所提出模型的有效性,本文在UBFC-rPPG和UBFC-phys数据集上进行验证。首先,对模型参数进行配置;其次,其他对比模型根据文中的参数设置,在rPPG-Toolbox平台上实现,并在UBFC-rPPG数据集上进行训练,在UBFC-phys数据集上进行测试。同时,FVSC-HR模型在UBFC-rPPG数据集上进行训练,在UBFC-phys数据集上进行测试,最终得到表2所示交叉验证下不同模型的性能统计结果。

Table 2 Performance statistics of different models on UBFC-phys dataset

方法	年份	MAE	RMSE	MAPE
DeepPhys ^[20]	2018	30.1	30.1	26.8
PhysFormer ^[8]	2022	13.5	15.9	12.5
PhysNet ^[21]	2019	21.1	26.3	23.8
TS-CAN ^[22]	2020	12.0	12.4	13.7
TransPhys ^[10]	2023	5.1	10.2	-
FVSC-HR(Ours)	2024	7.3	7.4	7.2

“-”代表着原始论文中无相关评价指标的数据

根据表2中的实验数据,从MAE评价指标分析,DeepPhys模型的MAE值为30.1,PhysFormer模型的MAE值为13.5,PhysNet模型的MAE值为21.1,TS-CAN模型的MAE值为12.0,TransPhys模型的MAE值为5.1。本文提出的FVSC-HR模型的MAE值为7.3,在所对比的模型中位列第二。从RMSE评价指标分析,DeepPhys模型的RMSE值为30.1,PhysFormer模型的RMSE值为15.9,PhysNet模型的RMSE值为26.3,TS-CAN模型的RMSE值为12.4,TransPhys模型的RMSE值为10.2。FVSC-HR模型的RMSE值为7.4,在所对比模型中数值最低,性能达到最佳。从MAPE评价指标分析,DeepPhys模型的MAPE值为26.8,PhysFormer模型的MAPE值为12.5,PhysNet模型的MAPE值为23.8,TS-CAN模型的MAPE值为13.7。FVSC-HR模

型的 MAPE 值为 7.2, 在所对比模型中数值最低, 性能达到最佳。

总体而言, 本文提出的 FVSC-HR 模型在 MAE、RMSE 和 MAPE 评价指标上分别取得了 7.3、7.4 和 7.2 的性能表现。其中, MAE 指标在所有对比模型中位列第二, 而 RMSE 和 MAPE 指标均达到所有模型中的最优值, 展现了最佳性能。FVSC-HR 模型能够与代表性算法竞争, 主要得益于卷积门控线性单元的细粒化特征能力和时间差分多头注意力机制获取视频帧上下文信息能力。该研究成果可广泛应用到智能健康监测、远程医疗等领域, 为远程心率估计提供了解决方案。

3 结语

远程光电容积脉搏波描记法通过收集对人体皮肤表面脉冲引起的细微颜色变化来估算心率。然而, 基于远程的心率检测方法在提取信号时忽略了其时间维度上的特征信息。本文通过融入卷积门控单元获取相邻帧的细粒度特征, 设计基于时间差分卷积的多头注意力机制, 进而提出了基于深度可分离卷积的人脸视频心率检测方法。在 UBFC-rPPG 和 UBFC-phys 数据集上的对比实验表明, 本文方法在性能上优于所对比的 6 种方法。考虑到当前测试仅基于公开数据集, 未来在实际场景中使用真实心率监测数据验证算法性能是一个重要的研究方向。此外, 开发基于 FVSC-HR 模型的智能心率监测与健康分析系统, 并开展实际应用测试, 也是未来值得深入探索的工作。

参考文献:

- [1] XIAO H G, LIU T Q, SUN Y S, et al. Remote photoplethysmography for heart rate measurement: A review[J/OL]. Biomedical Signal Processing and Control, 2023-10-21. <https://www.sciencedirect.com/science/article/pii/S1746809423010418>.
- [2] VERKRUYSE W, SVAASAND L O, NELSON J S. Remote plethysmographic imaging using ambient light[J]. Optics Express, 2008, 16(26): 21434-21445.
- [3] DE HAAN G, JEANNE V. Robust pulse rate from chrominance-based rPPG[J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2878-2886.
- [4] WANG W, DEN BRINKER A C, STUIJK S, et al. Algorithmic principles of remote PPG[J]. IEEE Transactions on Biomedical Engineering, 2017, 64(7): 1479-1491.
- [5] PILZ C S, ZAUNSEDER S, KRAJEWSKI J, et al. Local group invariance for heart rate estimation from face videos in the wild[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: 1335-13358.
- [6] HUANG H, XIE H Y, HUANG C C. Logical multitask contactless heart rate estimation based on face video[J]. Computer Engineering and Design, 2023, 44(10): 3179-3185.
黄海, 谢昊岩, 黄驰程. 基于人脸视频的逻辑多任务非接触心率估计[J]. 计算机工程与设计, 2023, 44(10): 3179-3185.
- [7] WANG Y, RONG Z, WANG S Y. Non-contact heart rate measurement method based on 3D residual attention network[J]. Foreign Electronic Measurement Technology, 2023, 42(2): 53-59.
王宇, 戎舟, 王苏煜. 基于 3D 残差注意力网络的非触式心率测量方法[J]. 国外电子测量技术, 2023, 42(2): 53-59.
- [8] YU Z, SHEN Y, SHI J G, et al. PhysFormer: facial video-based physiological measurement with temporal difference transformer[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 4176-4186.
- [9] YU Z, SHEN Y, SHI J G, et al. PhysFormer++: facial video-based physiological measurement with SlowFast temporal difference transformer[J]. International Journal of Computer Vision, 2023, 131(6): 1307-1330.
- [10] WANG R X, SUN H M, HAO R R, et al. TransPhys: Transformer-based unsupervised contrastive learning for remote heart rate measurement[J/OL]. Biomedical Signal Processing and Control, 2023-06-08. <https://www.sciencedirect.com/science/article/abs/pii/S1746809423004913>.
- [11] LIU X, ZHANG Y T, YU Z T, et al. rPPG-MAE: self-supervised pre-training with masked autoencoders for remote physiological measurements[J/OL]. IEEE Transactions on Multimedia, 2024-02-09. <https://ieeexplore.ieee.org/document/10430462>.
- [12] SHI D. TransNeXt: robust foveal visual perception for vision Transformers[DB/OL]. <http://arxiv.org/abs/2311.17132>.
- [13] BOBBIA S, MACWAN R, BENEZETH Y, et al. Unsupervised skin tissue segmentation for remote photoplethysmography[J]. Pattern Recognition Letters, 2019, 124: 82-90.
- [14] SABOUR R M, BENEZETH Y, DE OLIVEIRA P, et al. UBFC-Phys: a multimodal database for psychophysiological studies of social stress[J]. IEEE Transactions on Affective Computing, 2023, 14(1): 622-636.
- [15] LIU X, NARAYANSWAMY G, PARUCHURI A, et al. rPPG-Toolbox: deep remote PPG toolbox[J/OL]. Advances in Neural Information Processing Systems, 2023-12-10. <https://dl.acm.org/doi/10.5555/3666122.3669117>.
- [16] LI B, ZHANG P P, PENG J Y, et al. Non-contact PPG signal and heart rate estimation with multi-hierarchical convolutional network[J]. Pattern Recognition, 2023, 139: 109421.
- [17] WANG W, WEI Z, YUAN J, et al. Non-contact heart rate estimation based on singular spectrum component reconstruction using low-rank matrix and autocorrelation[J/OL]. PLOS ONE, 2022-12-30. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0275544>.
- [18] JAISWAL K B, MEENPAL T. Heart rate estimation network from facial videos using spatiotemporal feature image[J/OL]. Computers in Biology and Medicine, 2022-11-09. <https://www.sciencedirect.com/science/article/pii/S0010482522010150>.
- [19] ZHANG X, YANG C, YIN R, et al. An end-to-end heart rate estimation scheme using divided space-time attention[J]. Neural Processing Letters, 2023, 55(3): 2661-2685.
- [20] CHEN W, MCDUFF D. Deepphys: video-based physiological measurement using convolutional attention networks[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 349-365.
- [21] YU Z, LI X, ZHAO G. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks[DB/OL]. <https://arxiv.org/abs/1905.02419>.
- [22] LIU X, FROMM J, PATEL S, et al. Multi-task temporal shift attention networks for on-device contactless vitals measurement[J]. Advances in Neural Information Processing Systems, 2020, 33: 19400-19411.

(责任编辑:陈维捷)