

基于视频信号的非接触式心率和血氧饱和度检测方法综述^①



赵 娅, 王世铎, 贾 迪

(东北石油大学 计算机与信息技术学院, 大庆 163318)
通信作者: 赵 娅, E-mail: 261980020273@mail.nepu.edu.cn

摘 要: 心率和血氧饱和度是反映人体健康状况极其重要的生理指标. 近年来, 基于成像式光电容积描记技术 (imaging photoplethysmography, IPPG) 的非接触式心率和血氧饱和度检测方法因为其方便快捷且受约束较少等优点开始逐步成为研究热点. 主要工作如下: 首先, 介绍了非接触式检测方法的背景和研究意义; 其次, 从目标区域检测和感兴趣区域 (region of interest, ROI) 选取两个方面总结并点明其研究现状以及未来改进方向; 再次, 从传统方法、信号处理结合深度学习方法和端到端方法 3 个方面对心率和血氧饱和度检测方法进行了总结, 并梳理了深度学习方法所使用的数据集以及在各个数据集中所展现的检测效果; 最后, 指出该领域所存在的亟待解决的问题以及未来的研究方向.

关键词: 成像式光电容积描记技术; 非接触式; 心率和血氧饱和度检测; 信号处理; 端到端

引用格式: 赵娅, 王世铎, 贾迪. 基于视频信号的非接触式心率和血氧饱和度检测方法综述. 计算机系统应用, 2024, 33(10): 26-36. <http://www.c-s-a.org.cn/1003-3254/9646.html>

Review on Non-contact Detection Methods of Heart Rate and SpO₂ Based on Video Signal

ZHAO Ya, WANG Shi-Duo, JIA Di

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: Heart rate and saturation of peripheral capillary oxygenation (SpO₂) are very important physiological indicators of human health. In recent years, non-contact heart rate and SpO₂ detection methods based on imaging photoplethysmography (IPPG) have gradually become a research focus as they are convenient and freely-applied. The main work is as follows. First, the study introduces the background and research significance of non-contact detection methods. Secondly, two aspects of target region detection and region of interest (ROI) are selected to summarize and clarify the research status and future improvement direction. Thirdly, the detection methods of heart rate and SpO₂ are summarized from three aspects: traditional method, signal processing combined with deep learning method and end-to-end method, and the data sets used in deep learning method and the detection effects displayed in each data set are sorted out. Finally, the paper points out the problems that need to be solved and the future research direction in this field.

Key words: imaging photoplethysmography (IPPG); non-contact; detection of heart rate and SpO₂; signal processing; end-to-end

心率 (heart rate, HR) 和血氧饱和度 (saturation of peripheral capillary oxygenation, SpO₂) 是反映人体健康

状况的重要指标, 近年来正在逐步引起人们的重视^[1-4]. 随着人们的健康意识逐步增强, 各种心率和血氧饱和

① 基金项目: 黑龙江省自然科学基金 (LH2022F006); 湖南省卫生健康委科研课题 (202103101352)

收稿时间: 2024-02-27; 修改时间: 2024-05-06; 采用时间: 2024-05-14; csa 在线出版时间: 2024-08-21

CNKI 网络首发时间: 2024-08-22

度检测方法正逐步进入人们的视线。目前,最为常用的检测方法为传统的接触式检测法,该方法因为较高的准确度被人们广泛使用。然而,在实际使用中,接触式检测方法的一些局限性给使用者造成了许多不便。例如在检测过程中需要受试者皮肤紧贴检测仪器并保持稳定,否则会导致检测精度大幅降低。该问题会导致受试者在使用仪器的过程中感到不方便和不舒适,并且手抖患者、婴幼儿、烧伤患者等都无法便捷地使用接触式仪器。

由于传统接触式检测仪器存在的一些不足,非接触式心率和血氧饱和度检测方法正在逐步引起人们的关注。该方法是运用朗伯-比尔定律,通过观察皮肤颜色的变化提取出光电容积描记信号(photoplethysmography, PPG),实现对心率和血氧饱和度等生理信号的检测。该方法解决了接触式检测方法存在的不方便、不舒适的问题,开始被人们广泛研究。

然而,非接触式心率和血氧饱和度检测方法依然有许多问题亟待解决:比如皮肤中所蕴含的生理信息极其微弱,极易受到光线变化等情况的干扰,使得心率和血氧饱和度信号参杂较多的噪声,导致精度较低。

本文从视频预处理、检测方法、数据集、评价指标以及未来研究方向5个方面对该领域的研究现状进行综述。其中,本文对深度学习方法在该领域的发展情况着重进行了梳理。

1 视频预处理相关研究

在提取生理信号之前需要对视频数据进行预处理,包括目标区域检测和ROI选取。因为学者们大多通过人脸视频进行心率和血氧饱和度检测,所以本文对众多人脸检测方法进行总结和梳理。

1.1 人脸检测方法研究

文献[5-9]选择使用Viola-Jones人脸检测算法^[10]对视频中的人脸进行识别。该方法使用Haar特征值进行目标检测,可以满足人脸实时检测^[11]。然而,将Viola-Jones人脸检测算法应用在非接触式心率和血氧饱和度任务中依然有许多不足。比如该算法只能大致识别人脸的位置,并用矩形框进行标识,会将许多背景信息混入其中,难以保障所提取生理信号的信噪比,当人脸出现晃动时该问题尤为严重。为解决上述问题,学者们在进行人脸检测的同时进行人脸追踪,在人脸区域生成相应坐标,通过追踪坐标的移动实现人脸定位。文

献[10]首先使用Viola-Jones人脸检测器对视频中的人脸进行初步检测并使用矩形框对人脸区域进行标识,然后使用判别响应映射拟合(discriminative response map fitting, DRMF)的方法找到人脸矩形框内的66个人脸坐标,之后使用Kanade-Lucas-Tomasi(KLT)算法逐帧对人脸进行跟踪。使用此方法可以有效地减少背景区域以及头部运动对检测结果造成的影响。

除了较为常用的Viola-Jones人脸检测算法,学者们还使用其他方法对人脸进行检测:文献[12]使用MICNN检测网络进行人脸检测,MTCNN模型是以深度学习算法为基础,使用3个CNN级联算法结构,将人脸检测和人脸特征点检测同时进行;文献[13,14]使用卷积神经网络(convolutional neural network, CNN)对人脸特征点进行检测,并且文献[14]在基础网络中添加了小滤波器,使得该方法在人脸检测方面取得了不错的精确度和鲁棒性;文献[15]为了解决人脸坐标在帧间漂移导致提取的生理信号混入较多噪声伪影的问题,选择使用SSD(single shot multibox detector)人脸检测器,同时,为了更好地区分人脸图像中的皮肤区域、非皮肤区域和背景区域,文献[15]采用了SEEDS(superpixels extraction via energy-driven sampling)图像分割算法,该算法的核心部分是使用快速梯度下降算法最小化能量函数,在保证精度的同时也具有较高的运行速度。

能否识别检测到清晰的人脸对后续工作起到至关重要的作用,它决定了能否在ROI中提取出信噪比较高的生理信号并影响最后的检测精度。选取准确的人脸检测和跟踪方法是学者们研究的重点问题。

1.2 ROI选取方法研究

人体皮肤所蕴含生理信息非常微弱且极易受到光照变换和运动的干扰,使得提取的生理信号中混入大量噪声最终导致检测精度的降低。

基于上述分析,学者们不断进行尝试和探索,以选取更加合适的ROI进行生理信息提取。ROI选取情况如表1所示。下面对选取情况进行详细介绍。

1.2.1 面部ROI选取情况

由于人的面部遮挡较少并且毛细血管分布较为密集,相较于其他部位可以提取出较为强烈的生理信号,因此,有众多学者选择将人脸作为ROI进行生理信号的提取。文献[9,16-18]选择将整张人脸作为ROI并提取生理信号,大量实验证实预测得出的心率和血氧饱和度数值并不稳定,精度有待进一步提升。

表1 ROI选取情况

方法	ROI个数	ROI位置
文献[9,16-18]	1	整张人脸
文献[5-7]	1	除额头之外的人脸区域
文献[8,19]	1	除双眼之外的脸颊区域
文献[20]	1	嘴唇
文献[21]	2	嘴唇+脖颈
文献[22]	4	额头+双颊+人脸中心点
文献[23-25]	1	手指
文献[26]	1	手掌

对人脸区域进行分析可以发现, 尽管额头分布较为密集的毛细血管, 但有时会被头发等物体遮挡, 导致提取得到的生理信号不够清晰, 造成预测精度的下降. 因此, 有学者开始对人脸区域进行裁剪, 不再使用整张人脸作为 ROI: 文献[5-8]对人脸进行裁剪, 选取整张人脸的 60% 区域, 将额头之外的部位作为 ROI; 在此基础上, 文献[8,19]考虑到受试者无意识的眨眼也会对预测结果造成干扰, 因此选取眼睛以下的人脸区域作为 ROI; 文献[20,21]认为嘴唇附近分布有密集的毛细血管并且遮挡较少, 是理想的 ROI; 不仅如此, 文献[21]还考虑到当人们进行呼吸时会导致嘴唇不自觉的运动, 导致从该区域提取的生理信号混入较多噪声伪影, 因此, 文献[21]选取脖子作为第 2 处 ROI, 将两处提取到的生理信号进行结合, 最大程度提高信噪比; 为了尽可能减少头部运动对检测结果造成的影响, 文献[22]选取额头、双颊以及人脸中间 4 处位置作为 ROI, 额头和双颊用以提取生理信号, 人脸中心区域与其他 3 个区域相连, 保证 ROI 随着人脸轻微移动而移动, 减少因头部运动而产生的伪影对检测结果造成的影响; 为了探究 ROI 的大小与提取血液容积波 (blood volume pulse, BVP) 信号信噪比之间的关系, 文献[27]分别选取 3×3 、 101×101 、 201×201 (单位: 像素) 3 个不同面积的 ROI 并计算这 3 个区域提取 BVP 信号的信噪比, 大量实验结果表明 BVP 信号的信噪比与 ROI 边长成平方关系, 与 ROI 面积成正比.

人脸部位具有较为密集的毛细血管并且大部分皮肤裸露在外, 能够方便地提取生理信息. 但是, 由于人脸区域分布众多较为灵活的肌肉, 会产生许多微表情, 使得提取的生理信号混入较多运动伪影, 导致检测精度降低; 并且, 人脸区域无意识的晃动和部分遮挡也会降低生理信号的质量. 如何解决上述问题, 依然是学者们持续研究的方向.

1.2.2 其他 ROI 选取情况

除了人脸区域, 学者们还将其他毛细血管分布密集部位选作 ROI, 并提取生理信号进行心率和血氧饱和度检测. 受到指夹式脉搏血氧仪检测原理的启发, 文献[23-25]选取手指末端作为 ROI 用以提取生理信号, 他们认为手指末端具有密集的毛细血管, 并且与人脸不同, 手指末端肌肉难以做出复杂的动作, 从而减少了运动伪影对检测结果的影响; 文献[26]分别选取人脸和手掌作为 ROI, 并进行实验比较这两个部位生理信号的信噪比, 实验结果表明, 由于人脸区域出现的不自觉的运动 (如眨眼、呼吸、吞咽等) 会产生较为强烈的运动伪影, 对提取到的生理信号造成较大程度的干扰, 而手掌区域不存在上述动作, 提取得到的生理信号的信噪比较为理想.

选取手指或者手掌部位作为 ROI 可以有效避免不自觉运动给检测精度带来的影响. 但是依然存在一些问题, 文献[28,29]经过实验证实, 当手部被挤压时会导致检测结果存在较大误差, 最终导致检测精度降低; 另外, 由于以人体其他部位作为 ROI 的数据集较少, 会导致检测模型泛化能力较弱, 难以将检测方法应用到实际生活中. 开发使用手指和手掌部位作为 ROI 进行心率和血氧饱和度检测的数据集是亟待解决的问题.

2 心率和血氧饱和度检测方法研究

在完成视频预处理之后, 接下来进行心率和血氧饱和度检测. 在该环节中, 首先进行 BVP 信号提取, 然后对其进行去噪处理以提高信噪比, 最后对心率和血氧饱和度进行预测. 本文将从检测方法选用的角度进行叙述.

2.1 传统方法

传统方法是指在 ROI 完成肤色变化信号提取之后, 使用信号处理方法和数字图像处理技术对该信号进行去噪从而转换为纯净的 BVP 信号, 并进行心率和血氧饱和度预测. 方法流程如图 1 所示.

在最初的研究中, 学者们^[30]并未考虑如何提高 BVP 信号的信噪比, 只是简单地提取 ROI 的像素值并进行平均, 然后通过傅里叶变换转换成频谱图并进行血氧饱和度的预测. 在光照保持稳定以及 ROI 保持静止的情况下可以保持较高的检测精度, 但是当光线强度发生变化以及 ROI 出现轻微移动时检测精度会大幅下降. 自此, 学者们开始研究如何在检测环境出现变化

以及 ROI 移动的情况下保证 BVP 信号的质量以及检测精度. 信号处理方法选用情况如表 2 所示. 下面对这些方法进行介绍.

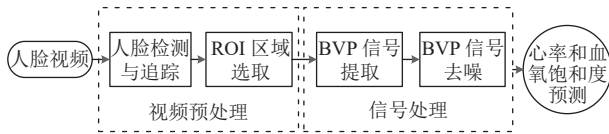


图1 传统方法检测流程图

表2 信号处理方法选用及检测内容

文献	方法	检测内容
文献[6]	ICA	心率
文献[17]	FastICA	心率
文献[31]	功率谱分析	心率
文献[32]	CEEMD+ICA	心率
文献[9]	CEEMD+ICA	血氧饱和度
文献[33]	WT+PCA+BSS	心率+血氧饱和度+呼吸率
文献[18]	SSA	心率

独立成分分析法 (independent component analysis, ICA) 是提取 BVP 信号较为常见的方法. 文献[6]考虑到血红蛋白在可见光和近红外光谱范围内的吸收率不同, 每个颜色传感器记录的是权重略有不同的原始信号的混合物. 并且受到盲源分离技术 (blind source separation, BSS) 的启发, 假设颜色传感器记录的源信号之间相互独立, 将 ICA 引入心率检测中来, 使得归一化的原始信号可以分解成 3 个独立的源信号, 并进行快速傅里叶变换 (fast Fourier transform, FFT), 得到较为纯净的 BVP 信号. 受到文献[6]的启发, 文献[17]使用 FastICA 算法得到最具有脉冲波特征的成分, 该方法可以只对所需成分进行分离, 所以文献[17]检测速度更快. 文献[31]通过在不同场景中对文献[6]中所提出的方法进行复现发现, 3 种成分在不同场景下都可以实现心率预测, 仅通过假设很难选择最优的成分. 为解决该问题, 文献[31]使用功率谱分析 (power spectral density, PSD), 将提取出的 9 个特征应用于 ICA 的输出分量, 其中 6 个特征包含最为强烈的频率, 并使用线性回归和 KNN (K-nearest neighbor, KNN) 算法对心率进行拟合.

除此之外, 学者们还尝试使用信号处理的其他方法并与 ICA 相结合以达到提取 BVP 信号提取和去噪的目的. 文献[32]将完备集合经验模态分解 (complete ensemble empirical mode decomposition, CEEMD) 技术

引入非接触式心率检测领域中, 在信号重建过程中对残余噪声进行精确衰减来提高其性能和效率, 然后选择具有最佳心肺频段的 IMF (intrinsic mode functions) 分量作为 ICA 的输入, 以去除运动伪影. 基于文献[34]的工作, 文献[9]将该方法应用在了非接触式血氧饱和度检测研究中. 由于非接触式血氧饱和度检测需要用到双通道比值法, 文献[9]在对原始信号进行通道分离后, 分别对 R 通道和 G 通道进行完备集合经验模态分解并生成不同的 IMF 分量, 然后使用来自红色通道的低频 IMF 和绿色通道的高频 IMF 作为 ICA 的输入, 实现信号去噪的目的. 文献[33]将小波变换 (wavelet transform, WT)、主成分分析 (principal component analysis, PCA) 和盲源分离 (BSS) 3 种信号处理方法相结合, 提出了 WT-PCA-BSS 算法, 实现了对心率、血氧饱和度和呼吸率的同时检测, 在光照环境发生变化时可以保持较高的准确度, 但是该方法中没有去除运动伪影的过程, 在头部出现大幅运动时难以保持原有的精度.

当受试者处于剧烈运动时, 上述方法无法有效去除运动伪影. 为去除运动伪影, 文献[18]将奇异谱分析技术 (singular spectrum analysis, SSA) 应用到 BVP 信号的去噪中. 该方法将 RGB 时间序列投影到 POS 平面和 CHROM 平面的交线上获取 BVP 信号并且还考虑了相邻帧间的面部区域中心的位移和相机的帧速, 获得了面部运动速度的时间序列信号, 并应用 SSA 提取了运动分量, 最后在脸部 ROI 的 BVP 信号中去除了 SSA 提取的运动伪影分量. 实验结果在自创数据集上显示出和真实值具有很高的相关性.

通过传统方法对心率和血氧饱和度进行预测往往需要凭借经验对受试者所处场景或者提取到的原始信号进行一定程度的假设, 但是仅凭经验很难对所有情况进行准确的判断, 使得当检测情景符合依据经验做出的假设时可以取得令人满意的结果, 然而当情景发生新的变化时检测精度或许会降低.

2.2 信号处理结合深度学习方法

随着研究的不断进步, 近年来, 学者们发现使用深度学习方法处理回归问题可以有效提高检测精度. 因此, 有学者开始尝试将信号处理方法与深度学习方法进行结合, 取得了优于传统方法的结果. 该方法的大致流程如图 2 所示. 由于视频预处理方法在第 1 节已经完成叙述, 此处不再重复. 此外, 目前尚未发现有学者

使用该方法对血氧饱和度进行检测,因此该处仅对心率检测方法进行叙述。

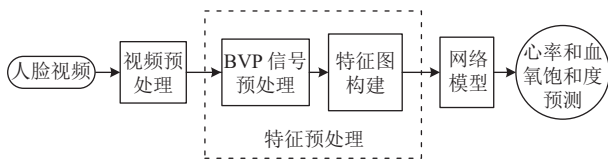


图2 信号处理结合深度学习检测方法流程图

2.2.1 特征预处理方法

在信号处理与深度学习结合的方法中,生成特征图的质量将直接关系到最终的检测结果,如何选取特征预处理方法以提高特征图的质量,是学者们的研究重点。

文献[34]首先将深度学习方法应用在了非接触式心率检测领域,在对颜色变化信息进行提取后,将二阶中心差、去趋势滤波和色度滤波3种方法进行结合对原始信号进行去噪,并采用快速傅里叶变换将信号转换成二维时频表示,而后传入网络进行回归运算。该方法有效地提高了检测精度,但是,由于需要使用一系列滤波方法对初始信号进行滤波,导致该方法较为复杂,时间复杂度较高。为解决该问题,文献[35]采用空间分解和时间滤波的方式直接从视频中提取包含心率信息相关信号的特征图像,避免了对原始信号进行滤波处理,很大程度提高了运算效率,可以做到对心率的实时预测。

在深度学习任务中,能否使用更好的方式对特征进行表示将在很大程度上对最终结果造成影响。文献[36]使用时空图对视频序列进行表征,使用该方法可以同时兼顾时间和空间的特征,使得后续网络可以学习到更多的信息。受到文献[36]的启发,文献[37-40]分别在此基础上进行改进,提出了不同的创建时空图的方法:为了尽可能全面地提取面部肤色所含有与心率有关的特征,文献[37]将选取的人脸ROI分成 10×10 的小块并对每个块的像素通道进行平均,然后按照帧序列将各通道像素的平均值进行拼接以生成时空图;为了解决视频中人脸缺失的问题,文献[38]提出了一种高质量时空图生成机制,该方法对视频中人脸大面积缺失或者过于模糊的帧进行去除并使用临近帧进行替换,由于帧间时间间隔远小于心跳时间间隔,因此使用该方法不会对心率预测产生不利影响,是切实可行的;文献[39]提出了多尺度时空图构建方法,尝试将不同颜色通道进行组合,使用RGB_YUV和RGB_HSV通道构建时空图;文献[40]使用小波变换的方法生成时空图以解决头部不同幅度的运动造成人脸图像模糊的问题,但是当光线扰动和头部运动问题同时出现时使用该方法容易造成检测精度的下降。

2.2.2 网络模型选取

在完成特征图构建之后,需要使用神经网络对特征图进行学习以求出心率数值。为此,学者们同样进行了许多研究,网络模型选用情况如表3所示。

表3 网络模型选取情况

方法	骨干网络	输入内容	输出结果
文献[34]	VGG15	时频图	平均心率
文献[35]	MobileNet	使用空间分解和时间滤波重构后的特征图	平均心率+瞬时心率
文献[36]	具有门控循环单元(GRU)的深度回归网络模型	时空图	平均心率
文献[37]	ResNet18(添加空间软阈值)	时空图	瞬时心率
文献[38]	深度卷积神经网络	时空图	平均心率
文献[39]	U-Net(rPPG)(添加多头注意力模块)	时空图	平均心率
文献[40]	ResNet18	时空图	平均心率

2.3 端到端深度学习方法

随着深度学习在非接触式心率和血氧饱和度检测领域中不断发展,端到端检测方法逐渐进入人们的视线。所谓端到端是指从输入端到输出端全程不需要人工进行干预,特征提取与生理信息预测均由神经网络完成。大致流程如图3、图4所示。

2.3.1 多阶段检测方法

在端到端深度学习的检测方法中,学者们通常使

用多阶段检测方法对心率和血氧饱和度进行检测,即在完成视频预处理之后使用不同的网络模型相继完成特征提取和生理信息预测。下面对该方法的研究现状进行简要概述。

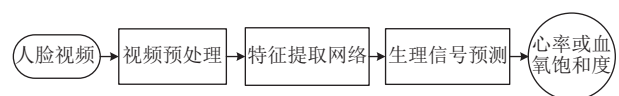


图3 多阶段端到端检测方法流程图

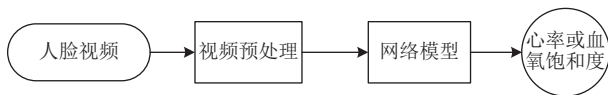


图4 单阶段端到端检测方法流程图

在视频中人脸颜色变化非常缓慢且极易受到头部运动和光照变化噪声带来的干扰。为解决该问题,文献[41]提出了一种新型的心率检测模型(ETA-rPPGNet),该模型分为时域段子网和主干网络:子网空间通过对视频进行分段,将长时间的视频内容转换为连续的局部信息并进行特征提取,然后使用平局池化函数对时间信息进行聚合,达到减少冗余信息的目的;主干网络通过添加时间注意力机制以时间信息学习能力,减少了噪声对心率预测的干扰。尽管使用三维神经网络对原始视频进行处理不对造成信息丢失,但是过大的模型参数会导致较高的时间复杂度。为解决网络预测模型运行时间过长的问题,文献[42]提出了一个“R(2+1)D”卷积模块对时空维度特征进行解耦,大幅减少了网络模型中的参数,提高了网络运行速度。并且,文献[42]还提出了基于皮肤的注意力机制,使网络自动忽略背景区域,有效降低非皮肤区域引入的噪声。由于目前没有较大规模的数据集,如何在数据规模有限的情况下完成网络模型的训练,实现非接触式心率检测是困扰学者们的一个关键问题。为解决该问题,文献[41]提出了FE-BE网络模型来进行心率预测。该模型由前件FE和后件BE组成,FE负责ROI选取以及信号提取和信号质量增强,BE负责心率预测。在该网络模型中,FE可以独立地在没有标注心率的数据集中进行训练从而可以自动对ROI进行检测并提取出高质量面部颜色变化信息并传入BE中进行心率预测,保证了心率检测精度和鲁棒性。

随着研究的逐步推进,学者们发现人脸视频中蕴含的一些底层特征同样对检测精度具有积极影响,将二者相互融合可以产生不错的效果。文献[43]提出了一种多特征融合的方法以实现血氧饱和度的检测。该方法首先使用残差与坐标注意力特征提取网络(residual and coordinate attention, RCA)对人脸信息进行提取,然后将提取的特征传入颜色通道模型(color channel model, CCM)进行信号重构从而充分利用人脸视频中的颜色特征信息。同时,该文献构建了基于网络的模型(network-based model, NBM)用于提取面部初始特征,并与CCM提取到的高级特征互相融合,达到提高检测

精度的目的。

2.3.2 单阶段检测方法

在端到端检测方法中,多阶段检测是目前较为主流的研究方向。然而,有学者开始考虑使用单阶段方法对心率进行检测,将特征提取网络和心率预测网络合二为一,可以更加快速地进行心率检测。文献[44]提出了X-iPPGNet单阶段心率检测网络模型。该模型以深度可分离卷积网络作为骨干网络,对每个颜色通道进行独立的时空卷积,其目的是提取输入图像序列每个颜色通道的局部特征,然后将各个通道的局部特征进行融合。此举可以更为全面地提取心率信息并有效地降低计算成本和内存需求。值得一提的是,针对单阶段检测方法的研究较少,文献[44]是目前唯一一篇提出使用单阶段网络模型进行非接触式心率预测的文章。

3 数据集

数据集对于研究至关重要。尤其是对于深度学习而言,数据集规模的大小、视频种类的多少都将直接关系到心率和血氧饱和度的检测精度以及网络模型的泛化性和鲁棒性。由于该领域研究的起始时间较晚,尚未出现较大规模的公共数据集,但是一些学者为了研究需要,按照自己的需求创建了一些数据集。本文按照公共数据集和自建数据集两种类型对部分数据集进行介绍。

3.1 公共数据集

MMSE-HR数据集^[45]:学者们构建该数据集的目的是在具有挑战性的环境下进行非接触式心率估计。其中包含102个RGB面部视频,以25 f/s的速度进行录制。受试者包括17名男性和23名女性,来自40个不同的民族。

MAHNOB-HCI数据集^[46]:该数据集是评估非接触式心率检测方法有效性、鲁棒性和泛化性的常用数据集。它包括来自12名男性和15名女性的527个视频以及相对应的生理信号。所有的视频都以61 f/s的速度录制,分辨率为780×580像素,利用心电信号计算真实的心率值。

UBFC-rPPG数据集^[47]:该数据集由42个受试者的42个视频组成。这些视频是用一个低成本的网络摄像头拍摄的,每秒30帧,分辨率为640×480像素,每次录制时间在50–90 s之间。使用Contec Medical CMS50E指夹式脉搏血氧仪记录心率值,并且与视频记录同步,

以建立真实的PPG信号。

BP4D+数据集^[45]: 该数据集是一个较大型的数据集, 它主要包括心率、呼吸频率和血压等生理参数。与现有心率数据集相比, BP4D+在数据量和种族多样性方面更具有优势。并且受试者在录制视频时做出更大幅度的运动。该数据库包含140名受试者以25 f/s的速度录制了1400个RGB颜色空间的视频, 每个视频时长为30 s–1 min, 分辨率为1040×1392像素。

VIPL-HR数据集^[48]: 该数据集是目前公开的专门用于心率和血氧饱和度检测的最大规模的数据集, 其中包含107个受试者总共录制了3130个视频。视频和生理信号数据是在受试者处在不同情景中采集的, 分别为静息状态、较大幅度运动状态、谈话状态、黑暗环境和光照环境等8种不同环境。

PURE数据集^[49]: 本数据集由10名受试者参与录制, 一共包含60段视频, 每个视频时长为60 s, 速率为30 f/s。数据集在不同的情境下采集, 分别为稳定状态、交谈状态、缓慢移动状态等6种状态。

OPF数据集^[50]: 该数据集用于研究远程生理信号测量, 包含200个5 min长的RGB视频, 由100名健康的成年人进行录制, 并提供相应的真实心电图信号。视频以60 f/s的速度录制, 分辨率为1920×2080, 压缩为MPEG-4格式。

LGI-PGGI数据集^[51]: 该数据集包括25名受试者, 总共录制了100段视频, 总时长大约200 min。视频和生理信号在4种不同情形中进行采集, 包括静息状态、头部略微运动状态、运动状态和谈话状态。视频由罗技HD C270进行录制, 帧率设置为25 f/s, 同步使用脉搏血氧仪进行心率检测。

DriverHeartRate数据集^[12]: 这是目前唯一的专门用于在行驶过程中对司机进行非接触式心率检测公开数据集。该数据集包含14名志愿者, 使用罗技C920r摄像头进行录制, 帧率为30 f/s, 录制场景在汽车内, 总共包括晴天、阴天、雨天、傍晚、夜间、雨夜等8种不同情景。在行驶过程中使用TM-204数字照度计对受试者面部接收到的照度进行检测。

COHFACE数据集^[52]: 该数据集由40名受试者参与录制, 总共包含160个视频。视频数据包含光照和自然两种状态。录制视频时同时采集BVP信号和呼吸信号。视频数据通过罗技HD C525进行采集, 分辨率为

640×480像素, 速率为每秒20帧。

ECG-Fitness数据集^[53]: 数据集中包含204个以健身为主题的视频, 使用两台罗技C920网络摄像头和FLIR热成像仪进行录制, 帧率为30 f/s, 分辨率为1920×1080, 以未压缩的YUV平面像素格式存储。受试者总共有17人。在录制过程中要求每名受试者进行4种不同类型的活动, 包括说话、划船、在固定自行车上锻炼以及在椭圆仪上锻炼。

3.2 自用数据集

PFF数据集^[54]: 该数据集由13名受试者参与录制, 摄像机使用尼康D5300, 帧率为50 Hz。录制场景分为明亮和黑暗两种类型, 受试者则分为轻微运动、静止和骑固定自行车3种状态。每段视频时长大约3 min, 分辨率为1280×720。

HR-D数据集^[55]: 该数据集包含75个视频, 平均时长82 s, 每秒22帧。视频的分辨率为800×480像素。为了使数据集更具有真实感和挑战性, 视频在录制过程中要求受试者处在不同方位、保持不同姿势, 不仅如此, 其中一些受试者在视频录制过程中还保留了胡须, 这给人脸检测和特征提取网络的设计提出了更高的要求, 可以促使心率检测模型更具有泛化性和鲁棒性。

FaceBio-v1数据集^[56]: 该数据集总共包含300个视频, 帧率为30 f/s, 来自300个年龄在18–26岁之间的实验对象。每个视频时长为1 min, 像素分辨率为1920×1080。这些视频是由荣耀v30在光线充足的环境下拍摄的。在录制过程中允许受试者变现出自然的面部表情和头部动作, 但是不允许受试者说话。在录制视频的同时使用BIOPAC MP160采集受试者的生理信号, 包括平均心率、呼吸频率、SpO₂、心电信号、血容量脉搏波。

4 评价指标

在研究过程中, 学者们使用不同评价指标来表示自己所提出的方法的有效性, 较为常用的评价指标为以下5种: 平均绝对误差(mean absolute error, MAE)、标准差误差(standard deviation, SD)、均方根误差(root mean squared error, RMSE)、平均错误率(mean error rate, MER)、皮尔逊相关系数(Pearson correlation coefficient, PCC)。表4对部分基于深度学习方法的评价指标以及所使用的数据集进行列举。

表4 不同估计方法性能比较

方法	数据集	MAE (bpm)	SD (bpm)	RMSE (bpm)	MER (%)	PCC
文献[35]	MAHNOB-HCI	—	4.65	4.27	4.95	0.79
文献[9]	MAHNOB-HCI	4.03	5.57	5.93	—	0.88
文献[15]	VIPL-HR	—	9.10	7.82	9.53	0.58
文献[44]	MAHNOB-HCI	2.08	3.47	3.41	2.73	0.92
	HR-D	3.02	7.14	6.58	3.43	0.81
文献[42]	COHFACE	4.67	—	6.65	—	0.77
	UBFC-RPPG	1.46	—	3.97	—	0.93
文献[56]	FaceBio-v1	1.75	2.95	3.43	—	0.93
	UBFC-RPPG	2.15	3.15	3.82	—	0.97
	COHFACE	5.57	7.69	9.50	—	0.75
文献[39]	VIPL-HR	4.09	—	5.88	—	0.93
文献[43]	UBFC-RPPG&&CONFACE	5.41	7.49	7.62	7.86	0.84
文献[46]	MMSE-HR	—	5.34	5.32	—	0.85
	UBFC-RPPG	4.99	6.25	6.26	—	0.67
	MAHNOB-HCI	3.17	3.93	3.93	—	0.88
文献[36]	MMSE-HR	—	6.95	6.95	6.55	0.98
	MAHNOB-HCI	—	2.79	3.26	3.67	0.95
文献[37]	MAHNOB-HCI	—	4.48	4.49	4.37	—
	MMSE-HR	—	6.86	6.83	6.21	—

5 总结及未来工作展望

基于视频信号的非接触式心率及血氧饱和度检测方法的研究,可以更好地满足人们尤其是婴幼儿、烧伤患者等一些特殊人群日常检测健康状况的需求。然而,学者们在该领域的研究依然存在易受噪声伪影干扰等问题,导致检测精度较低。为尽可能地提高检测精度,学者们在不同方面进行了诸多尝试。

首先,对视频信号进行预处理。在该环节中,学者们使用了包括 Viola-Jones、MTCNN 以及 SSD 在内的诸多人脸检测算法,在减少背景干扰方面取得了显著的效果。在提取到较为清晰的人脸之后,学者们将人脸不同部位选为 ROI 区域,尝试提取生理信号。大量实验与理论证明,人脸部位能够更加方便地提取较为强烈生理信息。但是如何减少人脸微表情以及无意识晃动对精度造成的干扰依然具有较大的研究空间。

在完成视频信号预处理之后便是对心率和血氧饱和度进行检测,其方法大致可以分为传统方法、信号处理结合深度学习以及端到端深度学习。在研究伊始,深度学习技术并未广泛普及,学者们大多采用传统方法进行检测,该方法通常需要凭借经验对受试者所处场景进行假设,导致其泛化性较低。随着深度学习技术的快速发展,人们开始采用信号处理结合深度学习的方法进行检测,其中较为经典的方法是将生理信号转换成时空图或时频图并传入神经网络进行特

征提取。该类方法有效提高了泛化性,但在生成时空图的过程中容易丢失部分生理信息,所以该方法依然有进步空间。为此,学者们使用端到端的方法进行检测,在检测过程中不需要人工干预,最大程度保留人脸中所蕴含的生理信息,并且检测方法的泛化性也得以进一步增强,该方法是目前使用频率较高的检测方法。

除此之外,本文还列举了该领域较为常用的数据集,以及文中的部分检测方法在公开数据集中所呈现的检测结果。

虽然基于视频信号的非接触式心率及血氧饱和度检测方法研究已经取得了较大进步,但依然存在一些问题和挑战。

第一,缺少大规模公开数据集。虽然 VIPL-HR 是目前最大规模的公开数据集,但是相较于深度学习而言规模依然较小,容易导致网络模型出现过拟合等问题。原因在于心率和血氧饱和度等生理信息涉及个人隐私,数据收集工作较难进行。除此之外,现有数据集中的视频都是在背景较为单一、受试人员运动状态比较规律、环境光线较为稳定的情况下进行的录制,难以模拟在真实环境下进行检测的情况。如何构建大规模公开数据集是学者们亟需解决的问题。

第二,更准确的 ROI 选取。就目前来看,学者们仅通过实验判定 ROI 的合理性,缺少从理论的角度分析选取 ROI 的原因。针对多种状态、多种情景、不同受

试者如何选取最优的 ROI 依然是需要深入研究的问题。

第三, 受试者隐私问题保护。人们在对自己的心率和血氧饱和度进行非接触式检测的过程中, 将频繁使用各类摄像头, 并且通过数据库保存检测结果, 存在用户人脸信息与生理信息泄露的风险。如何妥善保存视频与生理信息, 最大程度保护用户隐私是学者们在未来需要面临的重要挑战。

参考文献

- 1 Qi L, Yu HD, Xu LS, *et al.* Robust heart-rate estimation from facial videos using Project_ICA. *Physiological Measurement*, 2019, 40(8): 085007. [doi: [10.1088/1361-6579/ab2c9f](https://doi.org/10.1088/1361-6579/ab2c9f)]
- 2 Jain PK, Tiwari AK. Heart monitoring systems—A review. *Computers in Biology and Medicine*, 2014, 54: 1–13. [doi: [10.1016/j.combiomed.2014.08.014](https://doi.org/10.1016/j.combiomed.2014.08.014)]
- 3 Tulppo MP, Kiviniemi AM, Junttila MJ, *et al.* Home monitoring of heart rate as a predictor of imminent cardiovascular events. *Frontiers in Physiology*, 2019, 10: 341. [doi: [10.3389/fphys.2019.00341](https://doi.org/10.3389/fphys.2019.00341)]
- 4 Mather M, Thayer JF. How heart rate variability affects emotion regulation brain networks. *Current Opinion in Behavioral Sciences*, 2018, 19: 98–104. [doi: [10.1016/j.cobeha.2017.12.017](https://doi.org/10.1016/j.cobeha.2017.12.017)]
- 5 Poh MZ, McDuff DJ, Picard RW. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering*, 2011, 58(1): 7–11. [doi: [10.1109/TBME.2010.2086456](https://doi.org/10.1109/TBME.2010.2086456)]
- 6 Poh MZ, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 2010, 18(10): 10762–10774. [doi: [10.1364/OE.18.010762](https://doi.org/10.1364/OE.18.010762)]
- 7 周秦武, 隋芳芳, 白平, 等. 嵌入式无接触视频心率检测方法. *西安交通大学学报*, 2013, 47(12): 55–60. [doi: [10.7652/xjtub201312010](https://doi.org/10.7652/xjtub201312010)]
- 8 Li XB, Chen J, Zhao GY, *et al.* Remote heart rate measurement from face videos under realistic situations. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014. 4264–4271. [doi: [10.1109/CVPR.2014.543](https://doi.org/10.1109/CVPR.2014.543)]
- 9 Al-Naji A, Khalid GA, Mahdi JF, *et al.* Non-contact SpO₂ prediction system based on a digital camera. *Applied Sciences*, 2021, 11(9): 4255. [doi: [10.3390/app11094255](https://doi.org/10.3390/app11094255)]
- 10 Viola P, Jones M. Robust real-time object detection. *Proceedings of the 2nd International Workshop on Statistical and Computational Theories of Vision—Modeling, Learning, Computing, and Sampling*. Vancouver, 2001. 137–154.
- 11 贾海鹏, 张云泉, 袁良, 等. 基于 OpenCL 的 Viola-Jones 人脸检测算法性能优化研究. *计算机学报*, 2016, 39(9): 1775–1789. [doi: [10.11897/SP.J.1016.2016.01775](https://doi.org/10.11897/SP.J.1016.2016.01775)]
- 12 Wu BF, Chu YW, Huang PW, *et al.* Neural network based luminance variation resistant remote-photoplethysmography for driver's heart rate monitoring. *IEEE Access*, 2019, 7: 57210–57225. [doi: [10.1109/ACCESS.2019.2913664](https://doi.org/10.1109/ACCESS.2019.2913664)]
- 13 汪秀军, 曹大平, 曹宜, 等. 非接触式生理参数测量设备的研究. *信息技术*, 2018, 42(8): 17–22. [doi: [10.13274/j.cnki.hdzj.2018.08.005](https://doi.org/10.13274/j.cnki.hdzj.2018.08.005)]
- 14 郑银环, 王备战, 王嘉珺, 等. 深度卷积神经网络应用于人脸特征点检测研究. *计算机工程与应用*, 2019, 55(4): 173–178. [doi: [10.3778/j.issn.1002-8331.1710-0280](https://doi.org/10.3778/j.issn.1002-8331.1710-0280)]
- 15 Hsu GSJ, Xie RC, Ambikapathi A, *et al.* A deep learning framework for heart rate estimation from facial videos. *Neurocomputing*, 2020, 417: 155–166. [doi: [10.1016/j.neucom.2020.07.012](https://doi.org/10.1016/j.neucom.2020.07.012)]
- 16 Verkrusse W, Svaasand LO, Nelson JS. Remote plethysmographic imaging using ambient light. *Optics Express*, 2008, 16(26): 21434–21445. [doi: [10.1364/OE.16.021434](https://doi.org/10.1364/OE.16.021434)]
- 17 赵飞帆, 方路平, 陈仕骁. 基于人脸视频测量脉搏方法的初步研究. *生物医学工程学杂志*, 2012, 29(5): 876–879, 918.
- 18 Xie K, Fu CH, Liang HG, *et al.* Non-contact heart rate monitoring for intensive exercise based on singular spectrum analysis. *Proceedings of the 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. San Jose: IEEE, 2019. 228–233. [doi: [10.1109/MIPR.2019.00048](https://doi.org/10.1109/MIPR.2019.00048)]
- 19 Kong LQ, Zhao YJ, Dong LQ, *et al.* Non-contact detection of oxygen saturation based on visible light imaging device using ambient light. *Optics Express*, 2013, 21(15): 17464–17471. [doi: [10.1364/OE.21.017464](https://doi.org/10.1364/OE.21.017464)]
- 20 Shao DD, Liu CB, Tsow F, *et al.* Noncontact monitoring of blood oxygen saturation using camera and dual-wavelength imaging system. *IEEE Transactions on Biomedical Engineering*, 2016, 63(6): 1091–1098. [doi: [10.1109/TBME.2015.2481896](https://doi.org/10.1109/TBME.2015.2481896)]
- 21 Wei B, He X, Zhang C, *et al.* Non-contact, synchronous dynamic measurement of respiratory rate and heart rate based on dual sensitive regions. *BioMedical Engineering OnLine*, 2017, 16(1): 17. [doi: [10.1186/s12938-016-0300-0](https://doi.org/10.1186/s12938-016-0300-0)]
- 22 Casalino G, Castellano G, Zaza G. Evaluating the robustness

- of a contact-less mHealth solution for personal and remote monitoring of blood oxygen saturation. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(7): 8871–8880. [doi: [10.1007/s12652-021-03635-6](https://doi.org/10.1007/s12652-021-03635-6)]
- 23 Jonathan E, Leahy M. Investigating a smartphone imaging unit for photoplethysmography. *Physiological Measurement*, 2010, 31(11): N79–N83. [doi: [10.1088/0967-3334/31/11/N01](https://doi.org/10.1088/0967-3334/31/11/N01)]
- 24 Gregoski MJ, Mueller M, Vertegel A, *et al.* Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. *International Journal of Telemedicine and Applications*, 2012, 2012: 696324. [doi: [10.1155/2012/696324](https://doi.org/10.1155/2012/696324)]
- 25 Nam Y, Kong Y, Reyes B, *et al.* Monitoring of heart and breathing rates using dual cameras on a smartphone. *PLoS One*, 2016, 11(3): e0151013. [doi: [10.1371/journal.pone.0151013](https://doi.org/10.1371/journal.pone.0151013)]
- 26 Wu R, Gao HY, Zhang C, *et al.* To investigate the feasibility of palm-video-based IPPG technique. *Proceedings of the 5th International Conference on Information Communication and Signal Processing (ICICSP)*. Shenzhen: IEEE, 2022. 296–302.
- 27 曾旭, 代俊丽, 赵建, 等. IPPG 信号的噪声估计与分析. *航天医学与医学工程*, 2021, 34(1): 52–58. [doi: [10.16289/j.cnki.1002-0837.2021.01.009](https://doi.org/10.16289/j.cnki.1002-0837.2021.01.009)]
- 28 Sun Y, Hu S, Azorin-Peris V, *et al.* Noncontact imaging photoplethysmography to effectively access pulse rate variability. *Journal of Biomedical Optics*, 2012, 18(6): 061205. [doi: [10.1117/1.JBO.18.6.061205](https://doi.org/10.1117/1.JBO.18.6.061205)]
- 29 Lenskiy AA, Lee JS. Driver's eye blinking detection using novel color and texture segmentation algorithms. *International Journal of Control, Automation and Systems*, 2012, 10(2): 317–327. [doi: [10.1007/s12555-012-0212-0](https://doi.org/10.1007/s12555-012-0212-0)]
- 30 Wieringa FP, Mastik F, Van Der Steen AFW. Contactless multiple wavelength photoplethysmographic imaging: A first step toward “SpO₂ camera” technology. *Annals of Biomedical Engineering*, 2005, 33(8): 1034–1041. [doi: [10.1007/s10439-005-5763-2](https://doi.org/10.1007/s10439-005-5763-2)]
- 31 Monkarsi H, Calvo RA, Yan H. A machine learning approach to improve contactless heart rate monitoring using a webcam. *IEEE Journal of Biomedical and Health Informatics*, 2014, 18(4): 1153–1160. [doi: [10.1109/JBHI.2013.2291900](https://doi.org/10.1109/JBHI.2013.2291900)]
- 32 Al-Naji A, Perera AG, Chahl J. Remote measurement of cardiopulmonary signal using an unmanned aerial vehicle. *IOP Conference Series: Materials Science and Engineering*, 2018, 405: 012001. [doi: [10.1088/1757-899X/405/1/012001](https://doi.org/10.1088/1757-899X/405/1/012001)]
- 33 嵇晓强, 刘振瑶, 李炳霖, 等. 面部视频非接触式生理参数感知. *中国光学*, 2022, 15(2): 276–285. [doi: [10.37188/CO.2021-0157](https://doi.org/10.37188/CO.2021-0157)]
- 34 Hsu GS, Ambikapathi AM, Chen MS. Deep learning with time-frequency representation for pulse estimation from facial videos. *Proceedings of the 2017 IEEE International Joint Conference on Biometrics*. Denver: IEEE, 2017. 383–389. [doi: [10.1109/BTAS.2017.8272721](https://doi.org/10.1109/BTAS.2017.8272721)]
- 35 Qiu Y, Liu Y, Arteaga-Falconi J, *et al.* EVM-CNN: Real-time contactless heart rate estimation from facial video. *IEEE Transactions on Multimedia*, 2019, 21(7): 1778–1787. [doi: [10.1109/TMM.2018.2883866](https://doi.org/10.1109/TMM.2018.2883866)]
- 36 Niu XS, Han H, Shan SG, *et al.* SynRhythm: Learning a deep heart rate estimator from general to specific. *Proceedings of the 24th International Conference on Pattern Recognition*. Beijing: IEEE, 2018. 3580–3585. [doi: [10.1109/ICPR.2018.8546321](https://doi.org/10.1109/ICPR.2018.8546321)]
- 37 吕浩原. 基于深度学习的非接触式心率检测系统研究 [硕士学位论文]. 大庆: 东北石油大学, 2023.
- 38 Wu CL, Yuan ZY, Wan SH, *et al.* Anti-jamming heart rate estimation using a spatial-temporal fusion network. *Computer Vision and Image Understanding*, 2022, 216: 103327. [doi: [10.1016/j.cviu.2021.103327](https://doi.org/10.1016/j.cviu.2021.103327)]
- 39 Yu SN, Wang CS, Chang YP. Heart rate estimation from remote photoplethysmography based on light-weight U-Net and attention modules. *IEEE Access*, 2023, 11: 54058–54069. [doi: [10.1109/ACCESS.2023.3281898](https://doi.org/10.1109/ACCESS.2023.3281898)]
- 40 Jaiswal KB, Meenpal T. Heart rate estimation network from facial videos using spatiotemporal feature image. *Computers in Biology and Medicine*, 2022, 151: 106307. [doi: [10.1016/j.compbiomed.2022.106307](https://doi.org/10.1016/j.compbiomed.2022.106307)]
- 41 Hu M, Qian F, Guo D, *et al.* ETA-rPPGNet: Effective time-domain attention network for remote heart rate measurement. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 2506212. [doi: [10.1109/TIM.2021.3058983](https://doi.org/10.1109/TIM.2021.3058983)]
- 42 Yin RN, Jia RS, Cui Z, *et al.* PulseNet: A multitask learning network for remote heart rate estimation. *Knowledge-based Systems*, 2021, 239: 108048. [doi: [10.1016/j.knosys.2021.108048](https://doi.org/10.1016/j.knosys.2021.108048)]
- 43 Hu M, Wu X, Wang XH, *et al.* Contactless blood oxygen estimation from face videos: A multi-model fusion method based on deep learning. *Biomedical Signal Processing and Control*, 2023, 81: 104487. [doi: [10.1016/j.bspc.2022.104487](https://doi.org/10.1016/j.bspc.2022.104487)]
- 44 Ouzar Y, Djeldjli 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. *Computers in Biology and Medicine*, 2023, 154: 106592. [doi: [10.1016/j.combiomed.2023.106592](https://doi.org/10.1016/j.combiomed.2023.106592)]
- 45 Zhang Z, Girard JM, Wu Y, *et al.* Multimodal spontaneous emotion corpus for human behavior analysis. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 3438–3446. [doi: [10.1109/CVPR.2016.374](https://doi.org/10.1109/CVPR.2016.374)]
- 46 Soleymani M, Lichtenauer J, Pun T, *et al.* A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 2012, 3(1): 42–55. [doi: [10.1109/T-AFFC.2011.25](https://doi.org/10.1109/T-AFFC.2011.25)]
- 47 Bennett S, El Harake TN, Goubran R, *et al.* Adaptive eulerian video processing of thermal video: An experimental analysis. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(10): 2516–2524. [doi: [10.1109/TIM.2017.2684518](https://doi.org/10.1109/TIM.2017.2684518)]
- 48 Niu XS, Han H, Shan SG, *et al.* VIPL-HR: A multi-modal database for pulse estimation from less-constrained face video. *Proceedings of the 14th Asian Conference on Computer Vision*. Perth: Springer, 2019. 562–576. [doi: [10.1007/978-3-030-20873-8_36](https://doi.org/10.1007/978-3-030-20873-8_36)]
- 49 Stricker R, Müller S, Gross HM. Non-contact video-based pulse rate measurement on a mobile service robot. *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication*. Edinburgh: IEEE, 2014. 1056–1062. [doi: [10.1109/roman.2014.6926392](https://doi.org/10.1109/roman.2014.6926392)]
- 50 Li XB, Alikhani I, Shi JG, *et al.* The OBF database: A large face video database for remote physiological signal measurement and atrial fibrillation detection. *Proceedings of the 13th IEEE International Conference on Automatic Face & Gesture Recognition*. Xi'an: IEEE, 2018. 242–249. [doi: [10.1109/FG.2018.00043](https://doi.org/10.1109/FG.2018.00043)]
- 51 Pilz CS, Zaunseder S, Krajewski J, *et al.* Local group invariance for heart rate estimation from face videos in the wild. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Salt Lake City: IEEE, 2018. 1335–13358. [doi: [10.1109/CVPRW.2018.00172](https://doi.org/10.1109/CVPRW.2018.00172)]
- 52 Heusch G, Anjos A, Marcel S. A reproducible study on remote heart rate measurement. *arXiv:1709.00962*, 2017.
- 53 Spetlík R, Franc V, Cech J, *et al.* Visual heart rate estimation with convolutional neural network. *Proceedings of the 2018 British Machine Vision Conference*. Newcastle: BMVC, 2018. 1–12.
- 54 Yu ZT, Peng W, Li XB, *et al.* Remote heart rate measurement from highly compressed facial videos: An end-to-end deep learning solution with video enhancement. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul: IEEE, 2019. 151–160. [doi: [10.1109/ICCV.2019.00024](https://doi.org/10.1109/ICCV.2019.00024)]
- 55 Sabokrou M, Pourreza M, Li XB, *et al.* Deep-HR: Fast heart rate estimation from face video under realistic conditions. *Expert Systems with Applications*, 2021, 186: 115596. [doi: [10.1016/j.eswa.2021.115596](https://doi.org/10.1016/j.eswa.2021.115596)]
- 56 Li B, Zhang PP, Peng JY, *et al.* Non-contact PPG signal and heart rate estimation with multi-hierarchical convolutional network. *Pattern Recognition*, 2023, 139: 109421. [doi: [10.1016/j.patcog.2023.109421](https://doi.org/10.1016/j.patcog.2023.109421)]

(校对责编: 孙君艳)