arXiv:2302.03840v2 [cs.CV] 1 May 2023

MMPD：多领域移动视频生理数据集

Jiankai Tang[1], Kequan Chen[1], Yuntao Wang[1*], Yuanchun Shi[1],Shwetak Patel[2],Daniel McDuff[2],Xin Liu[2]
*通讯作者

摘要—远程光电容积描记法（rPPG）是一种颇具吸引力的方法，可便捷、无创地同步测量生理生命体征信号。近年来，公开基准数据集对该技术的发展及准确性的提升发挥了重要作用。然而，现有公开数据集仍存在不足。首先，尽管移动设备上的摄像机无处不在，专门使用手机摄像机录制的数据集却寥寥无几。其次，大多数数据集规模相对较小，因此在外观（如肤色）、行为（如运动）和环境（如光照条件）方面的多样性均有限。为推动该领域的发展，我们推出多领域移动视频生理数据集（MMPD），其中包含使用移动电话对 33 名受试者录制的 11 小时视频。该数据集旨在采集肤色、身体运动和光照条件方面具有更广泛代表性的视频。MMPD 内容全面，包含八种描述性标签，并可与 rPPG-toolbox [1]配合使用。主流无监督方法和神经网络方法验证了该数据集的可靠性。该数据集的 GitHub 仓库：
https://github.com/THU-CS-PI/MMPD_远程光电容积脉搏波描记法_数据集。

一、引言

远程光电容积描记法（rPPG）是一种光学技术，通过皮肤反射光的细微变化来测量心脏脉搏，即光电容积描记图（PPG） [2]。对心率等生命体征进行无干扰测量，是远程健康监测的一项关键技术，尤其有助于筛查和监测慢性心血管疾病患者。然而，传统医疗设备成本高、操作复杂，难以进行定期测量。尽管 rPPG 有诸多优势，现有基于视频的测量方法性能往往不稳定，且容易受到以下因素变化的影响：i）外观（例如肤色），ii）环境（例如光照），以及 iii）活动（例如身体运动的类型）。研究表明，由于反射光的信噪比较低，从肤色较深的人身上提取脉搏信号更加困难 [3]。光照变化会显著改变人脸的外观，使得检测血流引起的反射率细微变化更加困难 [4]。光线昏暗或过亮还会导致曝光不足或过度，并产生不必要的镜面反射，进一步遮蔽信号。视频中的运动伪影带来了严峻挑战；当前最先进的模型在人们运动时，难以生成精确的脉搏波形和心率。大多数模型尚未在用户进行说话或行走等自然活动时接受过广泛测试 [4], [5], [6]。

公开基准数据集是科学界极其宝贵的资源；然而，所有数据集都是有限的。对于远程光电容积脉搏波描记法（rPPG），现有数据集并未包含能够让研究人员针对上述所有维度（外观、环境和活动）系统测试模型的样本。例如，广泛使用的 UBFC-rPPG [7] 数据集主要包含静止受试者的视频，其 Fitzpatrick皮肤类型为 2–3。PURE [8]数据集包含的头部动作相对不自然，且采集对象主要也是 Fitzpatrick皮肤类型为 2–3 的受试者。最后，许多现有的公开 rPPG 数据集都是使用数码单反（DSLR）摄像机或专业影像公司的设备录制的。这与最普及的摄像机类型——智能手机相机——形成了对比。

为弥补现有公开远程光电容积脉搏波描记法数据集的不足，我们推出了多领域移动视频生理数据集（MMPD）。该数据集包含 33 名 Fitzpatrick皮肤类型为 3–6 的受试者、四种不同的光照条件（高亮度 LED、低亮度 LED、白炽灯、自然光）以及四种不同的活动（静止、头部旋转、说话和行走）。MMPD 中的所有视频均使用手机拍摄。本文的贡献如下：1）我们推出了 MMPD 数据集，这是首个包含不同肤色类型（Fitzpatrick 量表为 3–6）的受试者、多种光照条件以及各种真实场景运动情形的公开数据集。2）我们进行了全面的定量分析，以评估现有最先进的神经网络方法和无监督信号处理方法在该数据集上的表现。我们的目标是为研究人员提供一个数据集，以支持开发能够应对复杂、真实场景的算法，并解决基于相机的生理测量中的偏差问题。

II. 相关工作

目前有多个常用的远程光电容积脉搏波描记法数据集（PURE [8],MAHNOB-HCI [9],BP4D [10],VIPL-HR [11], COHFACE [12], UBFC-RPPG [7], MR-NIRP [13], VicarPPG- 2/CleanerPPG [14], Scamps [15]）。其中一些数据集专为远程光电容积脉搏波描记法而采集，另一些则用于一般生理学和计算机视觉研究。我们从这些优秀的数据集中选取了使用频率最高的三个数据集进行分析和比较。

UBFC [7]。UBFC-RPPG 数据集使用 Logitech C920 HD Pro 网络摄像头采集，分辨率为

[1] Jiankai Tang、KequanChen、Yuntao Wang、Yuanchun Shi 就职于中国清华大学yuntaowang@tsinghua.edu.cn。[2] Xin Liu、Shwetak Patel、Daniel McDuff 就职于美国华盛顿州西雅图的华盛顿大学{xliu0,shwetak,dmcduff}@cs.washington.edu。

表 I 数据集比较

数据集	帧	受试者	摄像机	传感器	肤色	运动	光照	运动
UBFC	57,420	42	Logitech C920 HD Pro	CMS50E	✗	✗	✓	✗
PURE	168,120	10	eco274CVGE	CMS50E	✗	✓	✗	✗
Scamps*	1,296,000	2800	/	/	✓	✓	✓	✗
MMPD	1,188,000	33	三星 Galaxy S22 Ultra	HKG-07C+	✓	✓	✓	✓

由于不同数据集包含时长各异的视频，因此这里以视频帧数量来计算大小。*Scamps是一个合成数据集，因此无法直接与其他数据集进行比较。

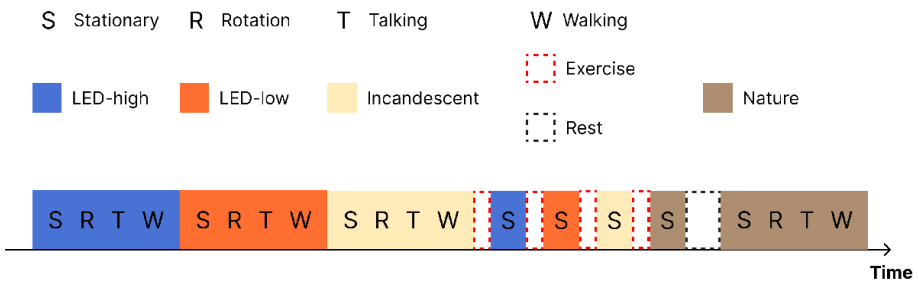

图1。数据采集协议的可视化示意图。在不同光照配置（蓝色）、活动（S、R、T、W）以及运动前后（红框与黑框）条件下，采集每位参与者的视频录制。

640x480，帧率为 30fps。使用 CMS50E 透射式脉搏血氧仪获取真实值 PPG数据，包括 PPG 波形和心率。受试者坐在距离摄像机约 1 米的位置，面部清晰可见；实验在室内进行，光照条件包括不同程度的阳光和室内光照。该数据集可靠且被广泛用作基准，但由于涵盖的肤色和动作范围有限，其多样性不足。

PURE [8]. PURE数据库包含60段时长一分钟的序列，这些序列使用eco274CVGE摄像机以30fps的帧率采集，分辨率为640x480像素；同时，通过CMS50E透射式脉搏血氧仪以60Hz的采样率并行采集PPG数据。该数据集因其动作多样性而被广泛使用，包括说话、平移和头部旋转，但在肤色、真实场景动作任务和光照条件方面缺乏多样性。

Scamps [15]。Scamps 数据集为 2,800 个视频序列提供了帧级真实值标签，涵盖光电容积描记图、心搏间期、呼吸波形、呼吸间期和 10 种面部动作。每段视频均根据相应的波形、动作单元强度以及随机采样的外观属性生成。尽管该数据集已展现出在各种应用中的潜力，但由于生命体征过于简单，在 SCAMPS 上训练的模型容易过拟合，因而往往表现不佳。

III. 数据集

为构建一个能够体现真实应用场景中视频多样性与复杂性的数据集，我们招募了来自不同国家的受试者，并在各种照明配置下进行了实验。共录制了 660 段时长为一分钟的视频，使用的设备是 SamsungGalaxy S22Ultra，同时使用 HKG-07C+ 血氧仪记录金标准PPG信号。本节将介绍数据采集协议、数据处理技术和数据集组织。

A. 数据采集

如前所述，光照和运动会显著影响从视频中提取光电容积脉搏波信号。为进一步研究这些因素，我们设计了一项同时采集面部视频和手指光电容积脉搏波信号的实验。实验流程如图 1 所示，所有视频均在允许触碰的距离下拍摄。

实验设置了四种光照强度和三类光源，包括低强度 LED 灯光（面部区域 100 流明）、中等强度白炽灯光（面部区域 200 流明）、高强度 LED 灯光（面部区域 300 流明）以及自然光（面部区域光照强度为 300 至 800 流明）。针对运动，我们设计了四项难度各异的任务：注视屏幕并保持静止、头部旋转、保持头部静止并说话，以及手持手机拍摄自拍视频。此外，我们还安排了四种运动，以研究身体活动对静止场景的影响。录制前，受试者需进行高抬膝或其他剧烈运动，使运动后心率升高。每项运动结束后，受试者都会充分休息，待平静下来后再进行下一项实验。

B. 数据处理

为了提高数据集的可访问性和易用性，我们对原始数据进行了预处理，并将其转换为便于使用的 MAT 文件格式，兼容 Matlab 和 Python。

表 II 无监督信号处理方法的结果。

方法	ICA [16]				POS [17]				CHROM [18]			
	平均绝对误差↓	均方根误差↓	平均绝对百分比误差↓	ρ ↑	平均绝对误差↓	均方根误差↓	平均绝对百分比误差 ↓	ρ ↑	平均绝对误差↓	均方根误差↓	平均绝对百分比误差 ↓	ρ ↑
肤色												
3	8.83	12.24	12.15	0.26	**5.76**	**9.67**	**8.63**	**0.48**	6.57	10.46	9.64	0.33
4	15.16	19.81	17.60	0.12	9.06	13.51	10.37	0.23	10.57	13.80	12.69	0.15
5	14.42	17.70	20.07	-0.10	12.78	16.69	19.29	-0.03	14.65	18.91	22.29	-0.12
6	17.14	21.52	19.77	-0.01	11.17	15.34	13.63	0.26	12.53	16.47	14.94	0.06
运动												
静止	11.48	15.82	15.06	0.16	9.70	13.74	13.35	0.26	10.23	14.23	14.46	0.15
旋转	11.75	15.88	15.35	0.06	7.50	11.99	10.85	0.40	9.28	14.02	13.10	0.16
说话	13.14	17.18	16.50	0.20	8.05	12.60	10.87	0.30	9.31	13.51	12.26	0.25
行走	26.15	30.75	27.43	-0.08	17.05	21.20	18.49	-0.06	17.61	21.07	19.15	-0.12
光照												
低亮度 LED	12.20	16.54	15.71	0.03	9.76	14.15	13.41	0.14	10.49	14.84	14.52	0.08
高亮度 LED	11.98	15.90	15.41	0.21	7.26	11.26	10.24	0.45	9.53	13.42	13.25	0.15
白炽灯	12.20	16.48	15.80	0.16	8.24	12.81	11.41	0.35	8.80	13.46	12.06	0.29
自然光	17.21	21.42	19.84	0.19	10.71	14.21	13.20	0.36	12.88	17.04	15.46	0.12

方法	GREEN [19]				LGI [20]				PBV [21]			
	平均绝对误差↓	均方根误差↓	平均绝对百分比误差↓	ρ ↑	平均绝对误差↓	均方根误差↓	平均绝对百分比误差 ↓	ρ ↑	平均绝对误差↓	均方根误差↓	平均绝对百分比误差 ↓	ρ ↑
肤色												
3	12.37	16.48	16.67	0.15	5.99	9.83	8.10	0.45	7.94	11.36	10.96	0.38
4	23.39	26.27	27.72	0.10	14.43	19.91	16.17	-0.17	15.87	20.50	18.34	-0.01
5	15.22	18.89	20.51	0.17	14.23	18.17	19.84	-0.02	14.62	17.77	20.17	0.08
6	20.59	24.96	23.37	0.13	17.02	22.15	19.28	0.03	17.24	21.38	19.81	0.10
运动												
静止	13.33	18.41	16.97	0.16	10.80	15.99	13.61	0.01	10.80	14.40	14.08	0.28
旋转	16.67	20.48	21.33	0.07	9.38	14.87	12.26	0.15	11.18	16.21	14.64	0.08
说话	17.16	21.14	21.48	0.06	11.36	16.26	13.87	0.20	13.44	17.43	16.69	0.19
行走	29.81	34.41	31.56	0.06	25.48	30.24	26.82	0.04	25.66	30.19	27.09	0.08
光照												
低亮度 LED	17.12	21.43	21.75	0.17	11.59	16.43	14.53	-0.01	11.91	16.09	15.16	0.11
高亮度 LED	14.71	19.13	18.76	0.08	9.74	14.91	12.46	0.17	13.01	17.22	16.95	0.09
白炽灯	15.33	19.49	19.27	0.06	10.22	15.78	12.75	0.19	10.49	14.76	13.31	0.33
自然光	20.07	24.74	23.19	-0.05	16.29	20.68	18.97	0.19	15.64	19.68	18.39	0.28

平均绝对误差 = 心率估计中的平均绝对误差（次/分钟），均方根误差 = 心率估计中的均方根误差（次/分钟），ρ = 心率估计中的皮尔逊相关系数。

视频的拍摄帧率为每秒 30 帧，分辨率为 1280x720 像素，但为了便于存储和传输，压缩至 320x240 像素。光电容积脉搏波信号从 200Hz 下采样至 30Hz，以匹配视频的帧率，因此每段视频包含 1800 帧。为了让研究人员能够探索各种因素对远程光电容积脉搏波描记法的潜在影响，我们为数据集添加了多个标签，例如肤色、性别、眼镜、头发覆盖情况和化妆。

为确保由不同设备采集的视频与真实值 PPG 波形同步，我们使用罗技 Yeti 麦克风作为中介。每次实验前，我们都会在两台设备上录制啁啾音频信号，然后计算两段录制音频信号之间的互相关，以确定手机与笔记本电脑之间的时间延迟。血氧仪的时间戳通过 USB COM 端口获取，从而利用两个时间戳同步光电容积脉搏波信号与视频信号。

C. 数据样本

图 2 展示了 MMPD 数据集中的一些样本。该数据集包含 Fitzpatrick 皮肤类型 3–6、四种不同的光照条件（高亮度 LED、低亮度 LED、白炽灯、自然光）以及四种不同的活动（静止、头部旋转、说话和行走）。

IV. 结果与讨论

A. 无监督信号处理方法 *s*

在我们的数据集 [19], [20], [17], [18], [16], [21] 上评估了六种传统无监督学习方法。在肤色比较中，我们排除了运动、自然光和行走条件，以消除任何混杂因素，并专注于当前任务。同样，运动比较实验排除了运动和自然光条件，而光照比较实验则排除了运动和行走条件。这种方法使我们能够排除混杂因素，更好地了解各项任务所面临的独特挑战。

表 III 使用 rPPG-toolbox [1] 在 MMPD 数据集上得到的基线结果。对于监督方法，我们展示了在 UBFC-rPPG 和 PURE 上训练的结果。

训练集	**UBFC** [7]				**PURE** [8]			
测试集	MMPD				MMPD			
	平均绝对误差↓	均方根误差↓	平均绝对百分比误差↓	ρ ↑	平均绝对误差↓	均方根误差↓	平均绝对百分比误差 ↓	ρ ↑
肤色								
3	3.60	6.91	5.01	0.76	**3.06**	**6.60**	**4.06**	**0.77**
4	14.45	20.51	16.23	-0.12	8.94	15.74	9.98	0.25
5	10.06	13.72	14.11	0.45	12.39	16.51	16.74	0.12
6	14.88	20.21	16.85	0.18	15.43	20.98	17.51	0.20
运动								
静止	5.34	11.17	6.32	0.56	5.91	11.56	7.13	0.54
旋转	11.73	16.45	15.14	0.12	8.92	14.99	11.24	0.17
说话	7.35	12.52	9.07	0.50	8.32	13.71	10.21	0.42
行走	24.91	29.76	26.24	-0.02	27.21	31.97	28.56	0.03
光照								
低亮度 LED	8.33	13.69	10.10	0.44	7.95	13.64	9.46	0.39
高亮度 LED	7.92	13.18	10.14	0.40	7.80	13.37	10.00	0.37
白炽灯	8.18	13.83	10.29	0.37	7.40	13.46	9.12	0.38
自然光	10.41	16.77	12.37	0.36	11.04	17.66	12.52	0.35

平均绝对误差 = 心率估计中的平均绝对误差（次/分钟），均方根误差 = 心率估计中的均方根误差（次/分钟），ρ = 心率估计中的皮尔逊相关系数。

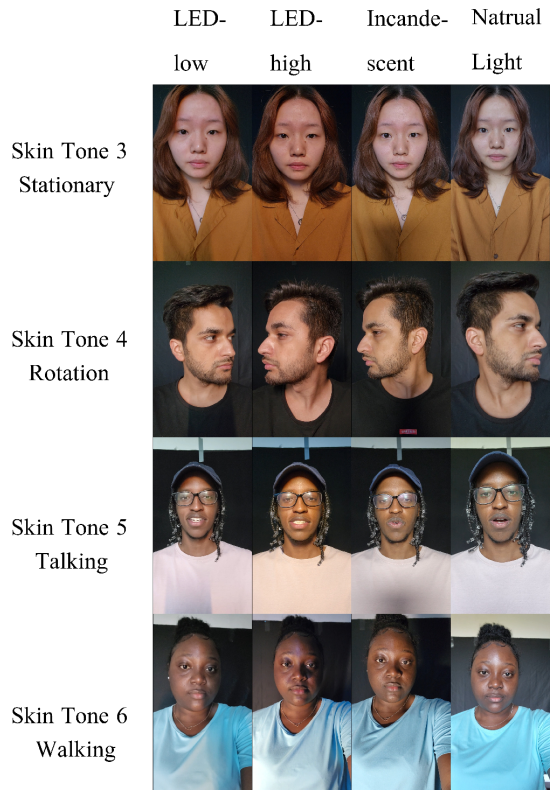

图 2. 涵盖肤色、动作和光照条件等多个领域的视频帧示例。

如表 2 所示，在无监督方法比较中，LGI [20] 方法在来自 3 型肤色的相对简单数据上表现最佳，而 POS [22] 方法在所有条件下的平均表现和稳健性方面均最佳。在肤色比较中，所有方法在 3 型肤色数据上都表现良好。然而，对于 4、5 和 6 型肤色，大多数结果显示，平均绝对误差（MAE）大于 10，表明泛化能力较差。在动作对比中，所有模型在难度最高的行走动作下表现都不佳，但各模型在静止、旋转和说话任务中各有所长。在光照对比中，三种人工光之间没有显著差异，且所有模型在自然光下表现都较差。

B. 监督式深度学习方法

本文还研究了先进的监督式神经网络在 MMPD 上的表现，并探讨了肤色、运动和光照的影响。我们使用了一个预训练的 TS-CAN [23] 模型，该模型在 UBFC [7] 和 PURE [8]数据集上训练。我们采用了与无监督方法评估相同的排除标准。

表 3 展示了监督式神经网络在不同任务中的结果。结果表明，该神经网络在所有场景下的泛化能力都不佳，因为它仅在 3 型肤色数据和静止任务中表现良好。这是因为训练数据（PURE 和 UBFC）仅包含 2–3 型肤色的受试者，且大多为静止视频。在 rPPG-toolbox [1]改进后的训练框架下，使用 UBFC[7]和 PURE[8]训练的模型之间没有显著差异。

C. 讨论

对研究发现的讨论表明，监督方法和无监督方法的性能会因测试数据与训练数据的相似程度而有所不同。我们的研究发现，当在4至6型皮肤的受试者身上进行测试，或面对具有挑战性的运动和光照条件时，监督方法的泛化能力有限。相反，无监督方法不依赖训练，因此表现出更强的泛化能力。我们的研究还发现，

公开的远程光电容积脉搏波描记法数据集可能无法充分反映 MMPD 数据集中遇到的现实挑战。具体而言，公开数据集对2至3型以外肤色类型的覆盖有限，视频大多为静止视频，且光照条件较为统一，因此监督方法的泛化能力有限。

为提高数据采集质量，我们建议在手机相机应用中使用原始模式，以捕捉面部的细微变化，并妥善放置手机。此外，尽量减少复杂的信号处理，并恰当使用 ffmpeg 等视频处理工具，有助于提高视频帧的质量，并缩短血氧仪与手机之间的时间延迟。鉴于手机视频中的人脸较大，还应考虑人脸对齐和帧填充。

总的来说，我们的研究强调，妥善选择训练数据和测试数据，并仔细考虑数据采集在真实场景中面临的挑战与局限性，对于提升远程光电容积脉搏波描记法方法的泛化能力和准确性至关重要。

V. 结论

本文介绍了 MMPD 数据集，该数据集由使用手机录制的 11 小时以上视频组成。数据集涵盖四种肤色、四种运动条件和四种光照条件下的受试者，为远程光电容积脉搏波描记法方法的基准测试提供了多样化的数据。MMPD 数据集包含八种描述性标签，旨在弥补现有手机录制数据集的不足，尤其是针对深色肤色人群的视频，以及真实场景中的运动和光照任务。MMPD 数据集及我们对远程光电容积脉搏波描记法方法的评估，推动了这项技术在准确性和泛化能力方面的发展，并有望改善医疗保健及其他应用。

致谢

本研究得到国家自然科学基金（NSFC）项目（批准号：62132010 和 62002198）、中国科协青年科学家资助计划项目（批准号：2021QNRC001）、清华大学自主科研计划、北京市网络多媒体重点实验室以及清华大学人工智能研究院的资助。本文所述涉及人体受试者的实验程序均已获得机构审查委员会批准。

参考文献

[1] Xin Liu、Xiaoyu Zhang、Girish Narayanswamy、Yuzhe Zhang、Yuntao Wang、Shwetak Patel 和 Daniel McDuff。深度生理感知工具箱。*arXiv* 预印本 *arXiv:2210.00716*，2022。[2] Daniel McDuff。基于摄像头的生理生命体征测量。*ACM* 计算调查 *(CSUR)*，2021。[3] Ewa M Nowara、Daniel McDuff 和 Ashok Veeraraghavan。肤色和性别对非接触式光电容积脉搏波描记法测量影响的元分析。载于*IEEE/CVF* 计算机视觉与模式识别会议研讨会论文集，第 284–285 页，2020。[4] Xin Liu、Yuntao Wang、Sinan Xie、Xiaoyu Zhang、Zixian Ma、Daniel McDuff 和 Shwetak Patel。Mobilephys：个性化的移动摄像头无接触生理感知。*arXiv* 预印本 *arXiv:2201.04039*，2022。

[5] Hao Lu、Hu Han 和 S Kevin Zhou。Dual-GAN：用于远程生理测量的血容量脉搏与噪声联合建模。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第 12404–12413 页，2021。[6] Daniel McDuff。通过深度超分辨率从视频中恢复生理信息。载于*IEEE*计算机视觉与模式识别会议研讨会论文集，第 1367 1374 页，2018 – 。[7] Serge Bobbia、Richard Macwan、Yannick Benezeth、Alamin Mansouri 和 Julien Dubois。用于远程光电容积描记法的无监督皮肤组织分割。《模式识别通讯》，124:82–90，2019。[8] Ronny Stricker、Steffen Müller 和 Horst-Michael Gross。移动服务机器人上的非接触式视频脉率测量。载于第 *23* 届 *IEEE*机器人与人类交互通信国际研讨会，第 1056–1062 页。IEEE，2014。[9] Mohammad Soleymani、Jeroen Lichtenauer、Thierry Pun 和 Maja Pantic。用于情感识别和隐式标注的多模态数据库。*IEEE*情感计算汇刊，3(1):42–55，2011。[10] Zheng Zhang、Jeff M Girard、Yue Wu、Xing Zhang、Peng Liu、Umur Ciftci、Shaun Canavan、Michael Reale、Andy Horowitz、Huiyuan Yang 等。用于人类行为分析的多模态自发情绪语料库。载于*IEEE*计算机视觉与模式识别会议论文集，第 3438–3446 页，2016。[11] Xuesong Niu、Hu Han、Shiguang Shan 和 Xilin Chen。VIPL-HR：用于从约束较少的面部视频中估计脉搏的多模态数据库。*arXiv* 预印本 *arXiv:1810.04927*，2018。[12] Guillaume Heusch、André Anjos 和 Sébastien Marcel。关于远程心率测量的可复现研究。*arXiv* 预印本 *arXiv:1709.00962*，2017。[13] Ewa Magdalena Nowara、Tim K. Marks、Hassan Mansour 和 Ashok Veeraraghavan。SparsePPG：探索利用近红外摄像头生命体征估计进行驾驶员监测。载于计算机视觉与模式识别（*CVPR*），第一届计算机视觉生理测量国际研讨会，2018。[14] Amogh Gudi、Marian Bittner 和 Jan van Gemert。利用干净的真实值进行评估的实时网络摄像头心率及心率变异性估计。应用科学，10(23):8630，2020。[15] Daniel McDuff、Miah Wander、Xin Liu、Brian L Hill、Javier Hernandez、Jonathan Lester 和 Tadas Baltrusaitis。Scamps：用于摄像头生理信号测量的合成数据。*arXiv* 预印本 *arXiv:2206 04197*，2022。.[16] Ming-Zher Poh、Daniel McDuff 和 Rosalind W Picard。使用网络摄像头进行非接触式多参数生理测量的进展。*IEEE* 生物医学工程汇刊，58 11，2010 1 7( ): – 。[17] Wenjin Wang、Albertus C den Brinker、Sander Stuijk 和 Gerard De Haan。远程PPG的算法原理。*IEEE*生物医学工程汇刊，64(7):1479–1491，2016。[18] Gerard De Haan 和 Vincent Jeanne。基于色度的rPPG实现稳健脉率测量。*IEEE* 生物医学工程汇刊，60(10):2878–2886，2013。[19] Wim Verkruysse、Lars O Svaasand 和 J Stuart Nelson。利用环境光进行远程容积描记成像。光学快报，16(26):21434–21445，2008。[20] Christian S Pilz、Sebastian Zaunseder、Jarek Krajewski 和 Vladimir Blazek。自然场景面部视频心率估计的局部群不变性。载于*IEEE*计算机视觉与模式识别会议研讨会论文集，第 1254–1262 页，2018。[21] Gerard De Haan 和 Arno Van Leest。利用血容量脉搏特征提升远程PPG的运动鲁棒性。《生理测量》，35(9):1913，2014。[22] Wenjin Wang、Albertus C den Brinker、Sander Stuijk 和 Gerard de Haan。远程PPG的算法原理。*IEEE*生物医学工程汇刊，64(7):1479–1491，2017。[23] Xin Liu、Josh Fromm、Shwetak Patel 和 Daniel McDuff。用于设备端无接触生命体征测量的多任务时序移位注意力网络。*NeurIPS*，2020。