

密 级：公开

学 号: 2210330136



Anhui Polytechnic University

硕士学位论文

题目

移动目标心率实时

监测方法研究

论文作者

程清扬

校内导师

王力超

校外导师

郝刚

专业学位类别

电子信息

研究方向 (领域)

检测技术与自动化装置

论文提交日期: 2024 年 5 月 31 日

类号: TP391

单位代码: 10363

密类级: 公开

学 号: 2210330136

题 目 移动目标心率实时监测方法研究

题 目 Research on Real-time Monitoring Method of
Moving Target Heart Rate

学 生 姓 名: 程清扬

校 内 导 师: 王力超

校 外 导 师: 郝刚

专业学位类别 : 电子信息

研究方向(领域): 检测技术与自动化装置

论文答辩日期: 2024 年 5 月 22 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：

程清扬

日期：2024 年 5 月 31 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 ____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：

程清扬

指导教师签名：

王机

日期： 2024 年 5 月 31 日

日期： 2024 年 5 月 31 日

移动目标心率实时监测方法研究

摘 要

心率是衡量个体生命活动健康状况的关键参数之一，在针对特定心血管疾病预防与治疗中具有不可忽视的生理学价值。因此，如何得到精准的人体心率测量值，对于提高人体健康水平具有十分重要的现实意义。现阶段，心率监测的方式包括接触式心率监测和非接触式心率监测。接触式心率监测最典型的方法是通过心电图信号来测定心率，然而这种方法在实施时要求使用者将湿润的电极片贴附在胸部，此举无疑给用户带来不便。非接触式的心率监测大多是在远程光电容积描记术的基础上进行研究，这种心率监测方法具有方便、安全、经济等优点。但该方法在实际运用时存在一定的局限性。真实场景下，目标处于移动状态，光照变化剧烈，采集的视频内存在冗余信息，这些因素都会一定程度的影响心率监测算法的鲁棒性。现阶段研究大多将对象设定在静止状态下，对于以上问题研究较少。因此，为了解决上述挑战，论文针对移动目标的心率实时监测方法展开研究，将移动目标作为研究对象，主要工作如下：

（1）针对移动目标场景，设计移动目标心率实时监测框架。首先完成基于视觉的心率监测框架搭建，对原始框架改进，添加移动目标检测模块，优化心率检测模块。改变不同的场景设定来验证改进后框架的有效性，最后得到对于移动目标有较高鲁棒性的监测框架。

（2）对于移动目标检测模型改进。选择 YOLOv4 为原始模型，复现 YOLOv4 模型。针对当前广泛使用的检测模型普遍存在参数规模庞大、难以在终端设备上有效部署问题，论文提出将原有

CSPDarkNet53 特征提取网络结构替换为更为精简高效的 GhostNet 网络,旨在有效削减模型复杂性及运算。在进行远距离人体目标检测时,面临的问题主要包括小目标识别度下降、因遮挡造成的检测精确度减低以及目标特征语义信息不充分等。为了解决这些问题,论文引入高效通道注意力机制,在主干网络提取有特征层端添加该模块,移动目标检测精度提升约 0.05%,而模型总体参数仅占原始模型的 17.7%。

(3) 真实场景下,成像器帧率和分辨率、受试者到成像器的距离、视频压缩等因素可能会影响通过 rPPG 方法得到的任何生理测量的准确性。低分辨率的视频包含的生理信息较少,难以提取到所需要的特征,并且极易受到环境因素的干扰,给心率的准确测量带来了较大的挑战。论文将对得到的视频进行预处理,把传统插值方法与基于学习的方法复现并进行比较,选择 Real-Esrgan 模型作为预处理步骤的方法,恢复丢失的面部生理信息。

(4) 研究运动干扰对于心率值的影响。被测目标移动,会引起的皮肤表面反射误差,当表皮反射误差累积到一定程度,生理信息误差将会变大,这会影响心率值的准确估计。为了解决这个问题,大部分研究者对整个脸部区域实施空间平均处理,去除表皮反射误差。然而经过实验可知,受到面部运动,皮肤光滑度等影响,脸部不同区域的生理信息质量是不同的,这个操作无法很好的应对挑战。于是在心率监测模块引入功率谱密度(PSD)指标得到面部各区域的频域权值,将各区域信号与其对应的权值融合,合并从面部各区域提取到的像素强度信息,最终将得到较为准确的心率值。将算法在移动目标数据集上与传统方法对比,心率测量精度提高了约 3.85%。

关键词: 非接触式测量, 心率监测, 感兴趣区域提取, 图像超分辨率, 移动目标检测

RESEARCH ON REAL-TIME MONITORING METHOD OF MOVING TARGET HEART RATE

ABSTRACT

Heart rate is one of the key parameters to measure the health status of individual life activities, and it has physiological value that can not be ignored in the prevention and treatment of specific cardiovascular diseases. Therefore, how to get accurate human heart rate measurement is of great practical significance for improving human health. At present, the methods of heart rate monitoring include contact heart rate monitoring and non-contact heart rate monitoring. The most typical method of contact heart rate monitoring is to measure the heart rate through ECG signals. However, this method requires the user to attach wet electrode pads to the chest when it is implemented, which undoubtedly brings inconvenience to the user. Non-contact heart rate monitoring is mostly based on remote photoelectric plethysmography, which has the advantages of convenience, safety and economy. However, this method has some limitations in practical application. In the real scene, the target is moving, the illumination changes dramatically, and there is redundant information in the collected video. These factors will affect the robustness of the heart rate monitoring algorithm to some extent. At present, most of the research sets the object in a static state, and there is little research on the above problems. Therefore, in order to solve the above challenges, this paper studies the real-time heart rate monitoring method of moving targets, taking moving targets as the research object, and the main work is

as follows:

(1) Aiming at the scene of moving target, a real-time monitoring framework of moving target heart rate is designed. Firstly, the heart rate monitoring framework based on vision is completed, the original framework is improved, and the moving target detection module is added to optimize the heart rate detection module. In the experiment, different scene settings are changed to verify the effectiveness of the improved framework, and finally a monitoring framework with high robustness to moving targets is obtained.

(2) Improvement of moving target detection model. Select YOLOv4 as the original model and reproduce the YOLOv4 detection model. In view of the problem that the widely used detection model generally has huge parameters and is difficult to be effectively deployed on the terminal equipment, this paper proposes to replace the original CSPDarkNet53 feature extraction network structure with a more streamlined and efficient GhostNet network structure, in order to effectively reduce the model complexity and computing requirements. In the long-distance human target detection, the problems mainly include the decrease of small target recognition, the decrease of detection accuracy caused by occlusion, and the insufficient semantic information of target features. In order to solve these problems, this paper introduces the efficient channel attention mechanism, and adds this module to the layer where the backbone network extracts features, which improves the accuracy of moving target detection by about 0.05%, while the overall parameters of the model only account for 17.7% of the original model.

(3) Improvement of detection model for long-distance moving targets. Select YOLOv4 as the original model and reproduce the YOLOv4 detection model. In view of the problem that the widely used detection

model generally has huge parameters and is difficult to be effectively deployed on the terminal equipment, this paper proposes to replace the original CSPDarkNet53 feature extraction network structure with a more streamlined and efficient GhostNet network structure, in order to effectively reduce the model complexity and computing requirements. In order to solve the problems of low attention of small-scale targets caused by long-distance detection of human targets, occlusion, low detection accuracy and insufficient semantic information of target features, this paper introduces an efficient channel attention mechanism and adds this module at the end of the main network to extract features, which improves the accuracy of moving target detection by about 0.05%, while the overall parameters of the model only account for 17.7% of the original model.

(4) To study the influence of exercise interference on heart rate. When the measured object moves, it will cause skin surface reflection error. When the skin reflection error accumulates to a certain extent, the physiological information error will become larger, which will affect the accurate estimation of heart rate. In order to solve this problem, most researchers carry out spatial average processing on the whole face area to remove the epidermal reflection error. However, through experiments, we know that the quality of physiological information in different areas of the face is different due to the influence of facial movement and skin smoothness, and this operation can not meet the challenge well. Therefore, the power spectral density(PSD) index is introduced into the heart rate monitoring module to get the frequency domain weights of each region of the face, and the signals of each region are fused with their corresponding weights, and the pixel intensity information extracted from each region of the face is merged, and finally a more accurate heart rate value will be

obtained. Compared with the traditional method, the accuracy of heart rate measurement is improved by about 3.85% on the moving target data set.

KEY WORDS: non-contact measurement, heart rate monitoring, region of inter extraction, image super-resolution, moving target detection

目录

摘 要	I
ABSTRACT	III
第 1 章 绪论	1
1.1 本章引言	1
1.2 研究背景和意义	1
1.3 国内外研究现状	3
1.3.1 非接触式心率监测算法研究现状	3
1.3.2 移动目标检测算法研究现状	6
1.3.3 超分辨率重建算法研究现状	7
1.4 主要研究内容与章节安排	8
1.4.1 主要研究内容	8
1.4.2 本文章节安排	9
1.5 本章小结	10
第 2 章 移动目标心率实时监测框架研究	11
2.1 本章引言	11
2.2 光电容积脉搏波描记法	11
2.2.1 光电容积脉搏波描记法原理	11
2.2.2 光电容积脉搏波描记法干扰因素分析	13
2.3 基于视觉非接触式心率监测算法框架	14
2.4 移动目标心率实时监测算法框架	15
2.5 本章小结	19
第 3 章 移动目标检测方法研究	20
3.1 本章引言	20
3.2 YOLOv4 算法原理	20

3.2.1 特征提取网络	21
3.2.2 空间金字塔池化	22
3.2.3 路径聚合网络	23
3.3 基于 GhostNet 的移动目标检测改进	24
3.3.1 GhostNet 模型	24
3.3.2 基于 GhostNet 的 YOLOv4 改进网络	26
3.4 基于高效通道注意力机制的移动目标检测改进	28
3.4.1 高效通道注意力机制	28
3.4.2 基于高效通道注意力机制的 YOLOv4 改进网络	29
3.5 实验过程与结果分析	30
3.5.1 数据集和实验环境	30
3.5.2 评价指标	31
3.5.3 实验结果分析	32
3.6 本章小结	34
第 4 章 基于 PSD 加权色度特征的心率估计方法研究	35
4.1 本章引言	35
4.2 基于输入优化的 rPPG 心率检测	35
4.2.1 基于传统超分辨率重建模型的输入优化	36
4.2.2 基于 Real-ESRGAN 模型的输入优化	37
4.3 基于 PSD 加权色度特征的脉搏波信号处理方法研究	40
4.3.1 Viola-Jones 算法	40
4.3.2 色度特征提取	43
4.3.3 信号加权与计算	44
4.4 基于输入优化的 rPPG 心率检测对比实验	47
4.5 基于 PSD 加权色度特征的心率估计验证实验	48
4.6 本章小结	49
第 5 章 实验与结果分析	51
5.1 本章引言	51
5.2 实验数据来源	51

5.3 实验评估指标选择	53
5.3.1 误差指标	53
5.3.2 Bland-Altman 图	54
5.4 实验结果分析	54
5.4.1 目标处于静止状态下的性能对比实验	54
5.4.2 目标处于运动状态下的性能对比实验	56
5.4.3 目标处于不同光照强度下的性能对比试验	57
5.4.4 一致性分析实验	58
5.5 本章小结	60
第 6 章 总结和展望	62
6.1 工作总结	62
6.2 展望	63
参考文献	65
攻读硕士期间发表学术成果	72
致谢	73

第 1 章 绪论

1.1 本章引言

本章首先概述了移动目标心率实时监测算法的研究背景和意义；接着分别详细介绍了视觉非接触式心率监测算法、目标检测算法、超分辨率重建算法的国内外研究现状。

1.2 研究背景和意义

心率可以反映一个人的健康状况，是重要生命体征之一，它代表着人体每分钟心跳的次数。心率在多个领域有着重要应用^[1-3]。而心血管疾病是威胁人类死亡的一个重要因素，有研究表明，心脏速率的波动与心脏健康状况紧密相连^[4-6]。若心率长期处于过高状态，比如每分钟超过 160 次，或是过低每分钟不足 40 次，可能会增加罹患心血管疾病的风险，严重情况还会发生心梗、猝死情况。因此，在预防心血管疾病以及治疗的过程中，心率数值是需要重点关注的一个生命指标。医疗机构会通过监测得到的人体心率值变化情况，来判断人体目前疾病的发展情况，同时也可以及时发现问题，及时实施医疗手段，降低心血管疾病存在的潜在风险。医院经常用到的心电图、电子血压脉搏仪、指压式脉搏血氧仪^[7-10]等都是接触式心率测量产品，部分产品如图 1-1、1-2、1-3 所示。这些检测方法应用较为广泛，并且测量较为精准，但仍存在一定的缺陷。这些接触式仪器不可避免的要与人体皮肤接触，然而有些被测者并不适合直接与仪器接触，并且当被测者与仪器接触时，可能会产生一些情绪波动，导致心率产生大幅波动，从而使得仪器测量得到的心率数值不准确，降低测量的精确度。并且一些特殊的仪器需要专业人士操作，仪器使用时需要耗费大量的人力物力，使用成本较高，这些仪器很难满足日常监测和预防需求。



图 1-1 脉搏血氧仪



图 1-2 心电图设备



图 1-3 小米手环

非接触式心率监测方法，它一定程度上解决了上述接触式心率测量方法存在的一些问题。越来越多的研究者将研究方向转向非接触式心率监测。非接触式心率监测大部分都是使用光电容积描记(Photo Plethysmo Graphy, PPG)^[11]来进行研究，该技术用来检测皮下微血管组织中的血液容积变化。简单的基于面部视频的非接触式心率监测^[12]只需要一个摄像头即可，实验较为方便。一些连接组织对光的吸收基本不变，如肌肉、骨骼、静脉等，但血液不同，吸收光自然也会因动脉中的血液流动而改变。摄像机采集到的视频包括皮肤反射信息，一部分镜面反射，不包括生理信息，另一部分是光穿透皮肤射到血管之后再反射出去的属于漫反射，其中包括生理信息，通过一系列的处理分析，将得到生理信息中存在的心率周期

性的变化规律。被测者的心率值就可以从心率周期性的变化规律中得到。非接触式心率监测方法采集的视频来源有所不同,有的来自面部皮肤有的则不是^[13],但是都是根据皮肤反射模型,再使用相关技术,例如信号去噪技术^[14]、特征提取技术^[15]等,得到较为准确的生理信息,最终得出被测者的心率值。基于面部视频的心率监测应用场景更为广泛,该方法具有操作简单实用性更广等优点,基于桡动脉视频进行心率监测^[16],存在一定局限性,受目标运动干扰较大,并且准确性较低。基于面部视频的心率监测研究时间较长,已经形成了较为成熟统一的监测框架。研究者将关注点更多地放在一些干扰因素^[17]对心率值带来的影响的分析上,寻找更加有效的方法解决这些挑战。经研究,环境光照变化,被测者运动干扰等因素影响较大。

然而,当前的大多研究场景局限于实验室内,被测者处于静止状态,在这些场景下研究得到的方法无法应用在医疗、体育等真实场景领域中,这些方法存在一定的局限性。针对移动目标场景下进行心率监测出现的精度低、实时性差等问题,研究并设计一个低成本,高精确度的心率监测方法是目前该领域主要研究的一个方向。

1.3 国内外研究现状

1.3.1 非接触式心率监测算法研究现状

近些年,越来越多的学者将目光放在非接触式心率监测领域。2020年,面对突发的新型冠状病毒^[18],如何降低医疗人员与受检者的直接接触,提高心率、体温等生理参数测量的准确性成为至关重要的问题。与此同时,针对病人和一般群体的生理参数的监测和管理也是一个巨大的市场需求。

在1938年,Hertzman等人使用一种基于PPG的脉冲信号采集方法,通过采集脉搏波的峰值信息,进一步提取出心率值。该方法选择人体为研究对象,使用仪器得到血容量脉搏(Blood Volume Pulse, BVP)的变化情况,并对其变化情况进行分析^[19],从人体不同部位(如手腕、手指等)提取出心跳、呼吸等生理信号。比如脉搏血氧仪就是通过仪器得到血管内血液容积的变化情况,从而对心率进行分析计算,但这种方式并不适用于一些特定的场景。

2008年,Verkrusse等人^[20]基于被测者面部视频进行研究,使用图像处理以及信号处理的方法提取视频中的生理信号,进而得到被测者的心率值,实现基于人脸视频的非接触式心率值的估计测量,并且通过大量实验验证了该方法的可行

性,说明了利用光电容积脉搏波描记方法估计心率值的有效性。为研究非接触式心率监测领域奠定了坚实的基础。然而,研究中仅使用了单纯的带通滤波操作对信号进行去噪,得到的信号中仍然存在着很多干扰信号,在复杂环境下采集到的人脸视频中,存在着大量由于目标移动、光照变化等因素带来的噪声信号,携带有心率的信号很容易被这些噪声掩盖。因此,单纯采用带通滤波已不能有效地过滤掉其中存在的噪声,无法精确地提取心率信号,得到准确的心率值。2010年,Poh等人^[21]同样也是基于面部视频进行心率研究,他们使用一般的摄像机采集面部视频,将三通道信号使用盲源信号分离的方法分解^[22],提出了一种高效的基于人脸视频心率监测方法。该算法首先将定位视频中的感兴趣区域,即面部部位,然后利用30秒长的滑动窗口对视频中的感兴趣区域检测,实现心率的连续监测。之后计算该区域的各通道像素强度,对于得到各区域像素均值使用独立分量分析(Independent Component Analysis, ICA)方法,得到被测者的生理信息并且进一步估计得到心率值。在此之后,许多研究者对这种方法进行了改进,即在ICA之前和之后增加了几个时域滤波器^[23]。最近的研究表明,在人脸图像中,同时对多个人脸感兴趣区域进行同步分析,可以有效提升对于BVP信号的提取效果。2013年,Haan^[24]等学者针对在运动场景下盲源分离方法存在的挑战,提出了一种基于色度特征的肤色模型(CHROM),该模型的思想是把色度线性组合成标准化肤色,即假定归一化的肤色在白色光照条件下都是一致的,这样模型可以适用于不同颜色的光照场景,同时该模型应用于目标移动场景下也有不错的表现。Wang等人^[25]在研究人体脉搏波的基础上,提出了2SR算法,该算法设定了一个皮肤颜色空间,该方法主要特点是更加关注色调的变化,跟踪色调变化情况来得到脉搏变化情况,提出了一个空间子空间旋转的概念。经过实验验证,2SR算法在被测者肤色有所差异的情况下,测量结果明显效果好于基于盲源分离的算法。之后,Wang等人^[26]的在之前的研究基础上进行改进,又进一步提出了POS算法,该算法特别之处在于将得到的面部视频划分为两类,一类是含有皮肤像素区域,一种是非皮肤像素区域,然后将感兴趣区域设定为带有皮肤像素的部分,将时间序列进行归一化操作,以此来得到脉搏波信号。而在2016年,Chwyl等人^[27]提出一种更高效的心率监测方法。首先,在面部感兴趣区域中选择合适的采样点,并对其进行追踪,之后使用红斑模型对采样点进行计算,然后借助贝叶斯估计原理的思想,得到PPG波形,将心率值看成 $[0.7,4.0]$ Hz频段内的最大峰值的频率。2018

年, Prakash 等人^[28]等人研究发现以往的算法被测者换成处于移动状态的被测者时, 得到的估计心率值与真实心率值相差较大, 于是引入有界卡尔曼滤波概念, 对移动中的目标进行跟踪, 从而来降低目标移动带来的干扰。

近几年, 由于深度学习的高速发展, 在 rPPG 领域使用深度学习(DL)方法已经出现, 并取得了良好的效果。2018 年, Spetlik 等人^[29]首次借助卷积神经网络构建了提取器(Extractor)以及心率测量器(HR estimator)两个组件, 首先将人脸视频序列输入信号采集器中, 得到 rPPG 信号, 再利用心率测量器对得到的 rPPG 信号进行计算, 最终得到被测者的心率值。Chen 和 McDuff^[30]在双流法的基础上提出了第一个端到端卷积神经网络 DeepPhys, 该网络使用归一化帧差作为输入, rPPG 信号作为输出, 由 Spetlik 提出的 HR-CNN, 使用受试者的面部图像序列作为输入, 心率作为输出。2019 年, Wang^[31]等人提出一种互补流, 即 rPPG 提取流的方法, 从面部区域提取可靠的 rPPG 信号。使用低秩约束引入特征提取流来引导网络学习特征, 提高了模型获得感兴趣区域的鲁棒性。对于 rPPG 信号提取流, 首先使用卷积神经网络提取 rPPG 信号, 然后使用两层 LSTM 网络对 rPPG 信号进行进一步细化。融合两个流之后得到特征并同时估计心率值。降低了噪声对心率监测的不利影响。然而, 真实场景下, 仍存在许多变量, 会对心率监测带来影响。Blackford^[32]曾指出成像器帧率和分辨率、受试者到成像器的距离、视频压缩、感兴趣区域等变量可能会影响通过 rPPG 方法得到的任何生理测量的准确性。所含信号质量较差的视频, 可利用的有用信号较少, 所需要的生理信号难以提取, 很难得到精确的心率值。

上述经典的算法都在心率监测领域有着不错的表现, 但她们的研究条件大部分都是将被测者设定在处于实验室的静止状态。而在利用 rPPG 技术提取所需生理信号进而得到所需生理指标时, 测量值的准确度都会受到人体刚性运动(呼吸, 眨眼等)、非刚性运动(走路, 跑步等)和光照变化等干扰因素的存在而受到影响^[33]。所以想要在一些算法中得到良好的实验结果, 只有使得被测者面部和环境背景的稳定性得到保证, 但是这不仅会给被测者带来不便, 而且算法的应用场景将会被大大的限制。

近些年, 在心率监测领域, 越来越多的研究者选择使用超分辨率重建算法来解决由摄像机、光源和皮肤带来的挑战。McDuff 等人展示了心率测量的准确性如何随视频质量的下降而下降, 并指出 rPPG 信号可以在相对较长的距离(高达

50m) 上通过图像超分辨率重建的手段一定程度的恢复, 有利于后续得到较为准确的心率值。ZitongYu^[34]等人首次尝试对抗视频压缩损失和从高度压缩的视频中恢复生理信号, 首先将具有 RGB 通道的人脸图像输入时空网络(STVEN+rPPGNet) 和 PhysNet^[35]来恢复信号, 以此来得到较为准确的心率值。而对于心率监测的研究, 人们逐渐将场景设定在非实验室的环境下, 以便更好的应用在现实生活中。因此, 开发能够在现实场景中稳定监测心率的 rPPG 方法具有重要的实用价值。

1.3.2 移动目标检测算法研究现状

在本文的研究框架中添加了移动目标检测模块, 以便于更好的定位被测目标及感兴趣区域, 一定程度的降低复杂背景对心率监测的影响。在本文中移动目标是处在移动状态下的行人, 而行人检测是目标检测中基于应用层面的需求延伸出的一个分支。行人检测的目的是从复杂环境下的采集到的视频内, 判断并选择人体目标, 将目标使用矩形框框出来, 以此得到所需要检测到的目标。目前, 行人检测在许多领域都有所应用, 该技术可以与行人跟踪, 行人重识别等技术结合, 应用于安防监控^[36]以及医疗^[37]等领域。该技术作为目标检测的分支之一, 一直以来备受各界关注。同时基于行人检测属于深度学习中重要且基础的一部分, 它的发展影响着深度学习领域中其他部分^[38-40]的发展。

早期研究者在进行行人检测方向的研究时, 采用的方法较为低效, 首先就是特征提取部分, 研究者需要自己选择需要提取的区域, 这就给模型添加了许多不确定性和随机性, 得到特征后需要将其分类, 同样也只能选择复杂的集成方法, 这大大的增加了模型的运行速度, 并且模型的精度也不能得到保证。这些复杂的方法大多都拥有相同的流程: 首先都要进行预处理步骤, 进行初次的去噪; 然后提取候选区域的特征, 并挑选出有代表性的特征向量, 最后将得到的特征信息进行反复训练, 之后得到的模型就可以实现行人检测的目的。

在 2005 年 Dalai^[41]等人提出了方向梯度直方图(Histogram of Oriented Gradient, HOG), 使用 HOG 做行人检测, 作者对行人检测所使用的特征集进行研究, 提出的 HOG 特征描述符比当时存在的特征集有更好的表现。与边缘方向直方图^[42]、SIFT 描述子^[43]、形状上下文^[44]相比, HOG 是在密集网格中进行的, 该特征是通过将小单元组成更大的块并归一化块得到的, 有效的提高了性能。作者用行人检测(行人基本上是直立的)测试集进行测试, 使用 SVM^[45]作为分类器, 采用滑动窗口的方式, 框出目标区域, 这样操作保证了模型的速度, 降低了它的复

杂度，经过改进后的算法更适用于检测道路交通中的行人。Felzenszwalb^[46]等人在 Dalai 等人的研究基础上进行扩展，提出了 DPM(Deformable Parts Model)算法，增强提取行人特征的能力。该算法是通用性检测模型，考虑到视角变化，光照变化，类内变化等因素影响，该算法将模型分解成多个子模型，而其中的子模型借助 HOG 的思想，得到组合模型的模板，然后将模板与目标匹配，进而实现对目标检测的目的。DPM 算法优于之前 HOG 等传统算法，适用范围更广，精度更高。但是该算法同样也存在一定的局限性，在检测过程中，模板是非常关键的一个因素，如果模板不合适，将会大大的影响到检测目标的效果。

近些年，深度学习不断发展，随着图像处理算法的提出，人们逐渐将其应用到各个领域。越来越多的研究者与时俱进，针对一些关键问题进行多角度的研究，并且得到较好的研究成果。行人检测领域也开始引入深度学习进行研究^[47]，本文非接触式心率检测框架为解决目标与相机距离较远导致的获取面部信息量不足的问题，引入基于深度学习的远距离移动目标检测模块。

基于深度学习的目标检测算法。针对特定问题，独立分析选择适合的基础模型结构，并且参考多种提高模型精度的算法，汲取重要思想，有选择性的与合适的模块结合，最终得到较为理想的模型。基于深度学习的目标检测算法通常按照阶段数分类，即两阶段检测和单阶段检测。其中，单阶段检测法包括 YOLO 系列网络^[48-51]与 SSD 网络^[52]。两阶段检测法包括：R-CNN^[53]、SPP-Net^[54]、Fast R-CNN^[55]和 Faster R-CNN^[56]等。深度学习因其具有多层网络，可以映射到任意函数，得到更丰富的特征信息，面对复杂问题，模型的稳定性较好，适用范围较广。

1.3.3 超分辨率重建算法研究现状

(1) 基于插值的超分辨率方法

基于插值的方法主要有最近邻插值法^[57]、双线性插值法^[58]和双三次插值法^[59]。它们同样都是在得到高分辨率图像过程中，运用到了插值函数^[60]，最近邻插值法首先确定已知的图像像素值，然后借此得到目标图像的像素值。双线性插法则是从不同位置的点为基础进行插值，而双三次插值法是以双线性插法为基础的改进后的方法。它们都属于线性插值算法，与之相对应的还有非线性插值算法。线性插值方法中，最近邻插值法速度较快运算量较小，但是效果较差。双三次插值法是三个算法中性能最佳的算法，但是计算量最大。

（2）基于重建的超分辨率方法

基于重建的超分辨率方法包括基于频域的方法和基于空域的方法。基于空域的经典的方法有很多，例如迭代反投影法^[61]、凸集投影法^[62]，最大后验概率法^[63]等。它们都拥有一定的先验信息，基于此对低分辨率的图像进行重建。但是它们对先验条件的依赖程度也有所不同，依赖性较强的算法有着灵活性较差的缺点。其中 POCS 法在空域法中表现最好，首先它对先验条件依赖性较弱，并且运算速度较快，得到高分辨的图像质量较好，有着不错的性能。

（3）基于机器学习的超分辨率方法

越来越多的超分辨率方法引入了机器学习的思想^[64-66]，大多数思路是从训练样本中学习到低分辨率图像与高分辨率图像之间的映射关系，然后根据这个关系重建得到目标图像。

2000 年，Freeman^[67]等人最先开始利用机器学习算法研究图像超分辨率重建，他们使用马尔可夫网络对这个世界建模，从样本中得到特征关系，然后利用得到的网络参数重新构建出所需要的高分辨率图像。2004 年，Chang 等人受到了流形组合方法的启发，提出了邻域嵌入方法^[68]。该方法利用了加权求和的思想，减弱了模型对于样本的依赖，但同时模型的灵活性也一定程度上的降低了。2012 年，Gao 等人对邻域嵌入算法进行改进并提出了稀疏邻域嵌入方法^[69]。该方法以字典学习和稀疏系数为核心，利用得到的系数重建，最终得到高分辨率图像。该方法相比于之前的模型，提升了精度，以及模型的灵活性。

1.4 主要研究内容与章节安排

1.4.1 主要研究内容

本文的研究内容是移动目标实时心率监测算法，主要分为以下几个方面：

（1）移动目标心率实时监测框架研究。研究基于视觉的接触式心率监测存在的不足与问题，通过提出移动目标心率实时监测框架在一定程度上解决问题，提高算法的鲁棒性。

（2）移动目标检测方法研究。对移动目标心率实时监测框架的第一个模块存在的问题，在原始模型上改进，提高检测的速度以及精度。

（3）对于输入视频的预处理研究。通过研究传统重建方法与基于机器学习的方法，选择并改进合适的模型，恢复视频中丢失的面部区域生理信号，有助于后续心率监测的进行。

(4) 基于 PSD 加权色度特征的脉搏波信号处理方法研究。研究移动目标对于心率监测的影响, 针对运动伪影, 研究并选择合适的脉搏波信号处理方法, 针对各面部区域信号不同提出基于 PSD 加权色度特征的脉搏波信号处理方法, 有助于解决运动伪影的干扰问题。

1.4.2 本文章节安排

本文主要章节安排如下:

第一章: 介绍了本课题的研究背景及意义。针对心率监测, 移动目标检测算法, 超分辨率重建算法三个领域的研究现状做出了详细的介绍和分析。介绍本课题的主要研究内容, 并介绍本文的各章节安排情况。

第二章: 研究基于视觉的心率监测框架, 对生理学基础进行分析, 解释光照吸收的光学原理, 详细分析光电容积脉搏波描记法的技术原理, 介绍技术流程, 分析了基于视觉的框架流程以及对心率监测结果的干扰因素以及出现的问题, 提出更适用于移动目标场景下的新框架。

第三章: 介绍移动目标心率监测算法的第一模块内容, 对于该模块选择的模型进行复现, 并发现应用于本文场景会出现参数量较大, 速度较慢, 精度降低等问题, 对原始模型进行改进, 提高模型运行速度, 提高检测精度, 利于后续心率监测的进行。最后使用经过筛选的 VOC2007 数据集与其他模型进行精度以及参数量方面的比较。

第四章: 本章将介绍移动目标心率监测算法框架的第二部分内容。首先, 针对传统算法应用于本文场景中出现的感兴趣区域包含生理信息较少, 目标运动出现运动伪影等问题进行详细研究。首先将研究成像器帧率和分辨率、受试者到成像器的距离、视频压缩、感兴趣区域等变量对于通过 rPPG 方法得到的生理测量值准确性的影响问题。通过对比选择合适超分辨率重建算法作为预处理步骤使用的算法。其次对于脉搏波信号处理方法进行研究, 提出对基于 PSD 加权色度特征的心率监测方法。最后将提出的方法在一些数据集中测试, 并且与其他传统方法进行性能比较。

第五章: 进行算法系统实验。使用公开数据集和自行采集的数据集验证, 与其他传统算法进行性能比较。进一步验证本文提出的方法在不同场景下的鲁棒性。

第六章: 该部分主要围绕论文的核心研究内容及实践成果进行了系统梳理和总结, 提炼出论文的亮点与创新点。指出当前研究存在的若干不足与局限性。在

此基础上，对未来的研究方向进行了前瞻性的规划，明确了有待进一步完善的方面，并有针对性地提出了相应的解决策略和后续研究思路。

1.5 本章小结

本章首先介绍本课题的研究背景及意义。针对心率监测，移动目标检测算法，超分辨率重建算法三个领域的研究现状做出了详细的介绍和分析。介绍本课题的主要研究内容，并介绍本文的各章节安排情况。

第 2 章 移动目标心率实时监测框架研究

2.1 本章引言

基于光电容积脉搏波描记法的非接触式监测方法适用性较为广泛，避免了仪器与人体皮肤的接触，适用于更多的场景。成像器采集到目标视频，通过信号处理，能够得到较为精准的心率值。它凭借着低成本便捷的优势，受到较多关注，有着广阔的应用场景。但是目前研究者的场景大多设定在实验室中，实验环境较为理想，当研究场景转换到真实世界时，就会出现感兴趣区域信号质量差导致可用信息较少，目标不易被捕捉等问题。得到的心率估计值与真实心率值之间相差较大，算法准确性较低。

本章主要介绍光电容积脉搏波描记法基本原理，基于视觉的心率监测框架流程，以及本文提出的针对移动目标出现的问题设计的非接触式心率监测框架，其中包括移动目标检测模块与心率估计模块。

2.2 光电容积脉搏波描记法

2.2.1 光电容积脉搏波描记法原理

光电容积脉搏波描记法利用透射或反射光的变化来检测生物组织血管中微循环血流量体积变化情况。由于具有无创，价格低廉等特点，在医疗和体育等领域得到了广泛的应用。血液中的血小板，直接胆红素，以及含氧血红蛋白等都能够吸收光线，而皮肤组织因其具有丰富的血管网络，从而对光线的吸收和反射更加敏感，更易被检测到。人体的血液容量变化规律可以一定程度上的反映人体心脏活动情况，血管内血液容量的每一次周期变化都会影响心跳周期性变化。这个技术正是利用这个规律，获取到反射光强的周期变化，从而推测出血管内血容量的周期变化。通过一些手段得到血容量变化规律后，再使用信号处理技术可以得到心跳的周期性变化规律，进而得到被测者的心率值。图 2-1 为采集到的 PPG 信号以及 ECG 信号示意图，可以发现两者具有相同的周期性变化。ECG 信号是使用仪器得到的准确心率信号，具有一定的参考作用，一般将其作为真实值进行参考。经过分析可知，光电容积脉搏波描记法得到的信号可以代表心率的周期性

变化情况，可用于心率测量，证明了该技术是可靠的。

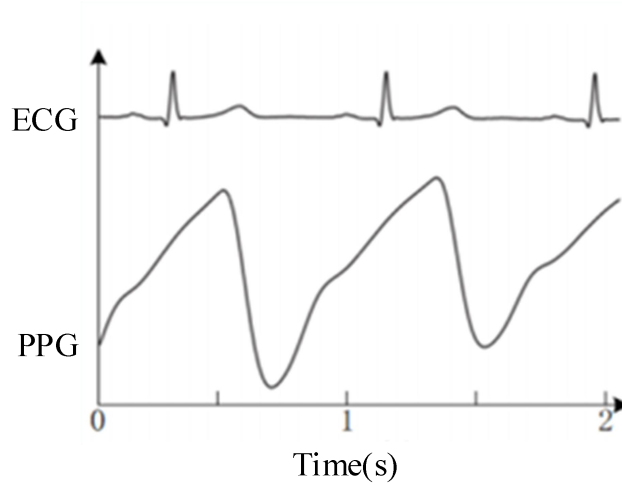


图 2-1 PPG 信号与心电图

光电容积脉搏波描记法原理图如图 2-2 所示。首先分析技术所需要的皮肤反射模型的建立过程。第一步需要假设被测者照射到的光线是来自同一个光源，并且光线照入皮肤，血管内的介质会改变光照的强度，即反射出的光照强度与照射进入的强度不同。那么就可以根据一些参数定义得到第 k 个皮肤像素的强度值：

$$C_k(t) = I(t)[v_s(t) + v_d(t)] + v_n(t) \quad (2-1)$$

其中， $C_k(t)$ 指第 k 个皮肤像素在 t 的像素强度值， $I(t)$ 表示光照强度，它由环境中光照颜色，亮度以及被测者距离光源距离等共同决定； $v_s(t)$ 表示镜面反射， $v(t)$ 表示漫反射，表示光照进入血管后经过介质之后反射出来的光； $v_n(t)$ 表示摄像机传感器的量化噪声。

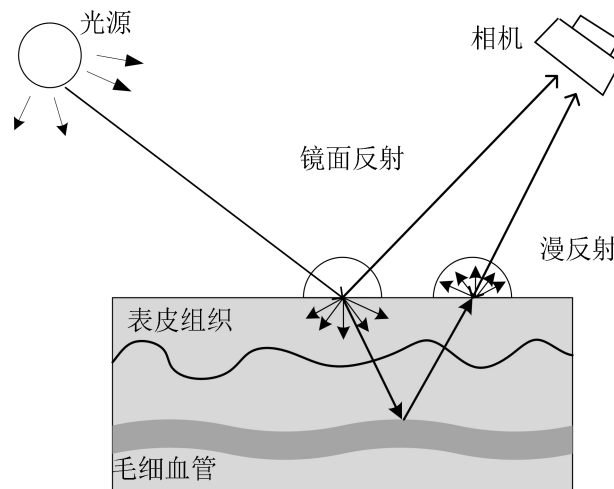


图 2-2 皮肤对光照的反射模型

光电容积脉搏波描记法借助光电技术记录血管内血容量的变化情况。假设光强度为 I ，波长为 λ 、入射光强度为 I_0 ，则透射光强度 I 可表示为：

$$I = I_0 e^{-\varepsilon(\lambda)AD} \quad (2-2)$$

式中， $\varepsilon(\lambda)$ 表示介质 ε 在特定波长 λ 下的吸收系数； A 是介质浓度； D 代表光线在该介质中穿过的距离。

在基于 rPPG 的心率测量方法中，固定光源照射到皮肤上，由于血管内介质浓度的改变，其反射与吸收的光强度也会随之改变，从而使血管内的血流体积也随之改变。另外，血管中的介质吸收光线的强度以及心跳的频率都会随着时间的推移而改变，因此，由照相机拍摄到的脸部光照反射的强度也会随着时间的推移而改变。由此，分析采集到的面部视频中反射出的光电信号就可以进一步得到与心率信号具有相似性的 rPPG 信号，然后再对信号进行处理进而得到被测者的心率估计值。

2.2.2 光电容积脉搏波描记法干扰因素分析

光电容积脉搏波描记法是一种利用成像装置等硬件设备，在可视条件下，将脉搏波信号与心率信号进行转换的一种非接触式的光学技术。所以，实际应用场景对光电容积脉搏波描记法的稳定运行影响很大，对于提高脉搏波信号质量以及获取心率测量值的准确性有极大的影响。在实际应用场景中，目标处于自然放松存在微小表情以及肢体运动状态。然而，光电容积脉搏波描记法很容易受到受试者状态，外部环境（光照影响等）以及成像设备等硬件设备的影响。同时，这些外在的影响也是无法避免的，以下将对影响光电容积脉搏波描记法的几个因素进行原因分析和原理介绍：

（1）环境光照：在实验室场景下，被测者接触到的光源相对稳定，真实场景下的光源更加复杂多变，从而导致真实场景下的心率监测具有一定的挑战性。外界环境的光线是光电容积脉搏波描记法依赖的条件之一，它决定了最终得到的心率值是否准确可靠。外界光线包括室内灯光以及室外自然光，都会间接的对心率监测造成干扰。光照条件不同，人脸图像的光谱分布同时也会随之发生变化，从而产生心率估测的误差。

（2）运动伪影：真实场景下，目标发生更多微小运动，会导致皮肤表面的

运动与变形,从而使得光线的散射和吸收发生变化,并且目标处于移动状态下会直接影响到血管内血液容积的变化情况,检测得到生理信号含有不少运动带来的噪声。而目标移动带来的噪声对心率信号的提取带来较大的挑战,在谱峰范围内运动伪影有时会与所要提取的心率信号存在部分重叠,而有些研究者采用简单的滤波方法来解决这个问题,但是这个操作无法消除运动伪影带来的影响,甚至存在把有用的心率信号当作噪声信号进行消除的可能,从而使得 rPPG 信号的信噪比大大降低。

(3) 成像设备: 成像系统的散焦, 成像设备与目标的相对运动, 以及设备自身固有的特质, 都可能会加大从原始信号中提取生理信号的难度。与成像器的距离同样也会影响最终获得的心率值。真实场景下, 来自成像器视频图像的面部区域通常很小, 分辨率较低, 难以准确定位人脸, 获得准确充足的脸部生理信息, 尤其大多是捕捉一个区域的整个场景, 与成像器的距离更是极大的影响了心率值的准确获得。

2.3 基于视觉非接触式心率监测算法框架

基于人脸视频的心率监测算法, 大多利用了光电容积脉搏波描记法。摄像头记录采集到人脸面部视频, 从视频中得到被测者皮肤包含的生理信息的变化情况, 进而在得到的周期性的生理信息变化情况中提取出相关的被测者的心率信息。

下面详细介绍基于视觉的非接触式心率监测算法框架, 流程图如图 2-3 所示。

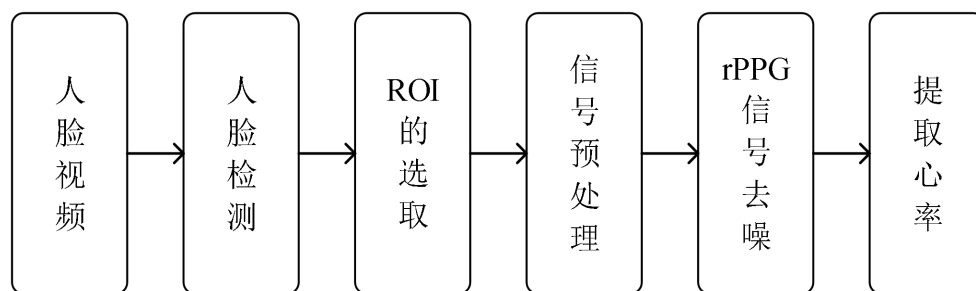


图 2-3 基于视觉的非接触式心率监测算法流程图

首先进行面部视频的采集, 通过成像设备进行。研究者使用手机等电子设备上的摄像头进行采集, 降低了基线漂移在此环节存在干扰的可能性。第二步对被测者的面部检测进行检测, 确定视频中人脸的区域。在提取到的人脸区域内进一步确定感兴趣区域, 感兴趣区域的准确选取关系着后续得到的生理信号的准确性, 之后提取到包含着噪声信号的原始生理信号。由于原始生理信号中携带较多由于

光照等干扰因素带来的噪声信号，需要对其进行预处理，大多数研究者都是将图像的像素的三通道分别提取出来，然后使用带通滤波等去噪手段去除噪声信号，得到较为纯净的生理信号。有的研究者会在这之后使用 ICA 等较为复杂的去噪方式进行二次去噪操作。对得到的较为纯粹的生理信号进行时频分析，从具有周期性的信号中得到被测者的心率值。另一种引入深度学习的心率检测框架则是在得到生理信号后使用神经网络训练针对心率检测的模型，从而得到被测者的心率值，如图 2-4 所示。

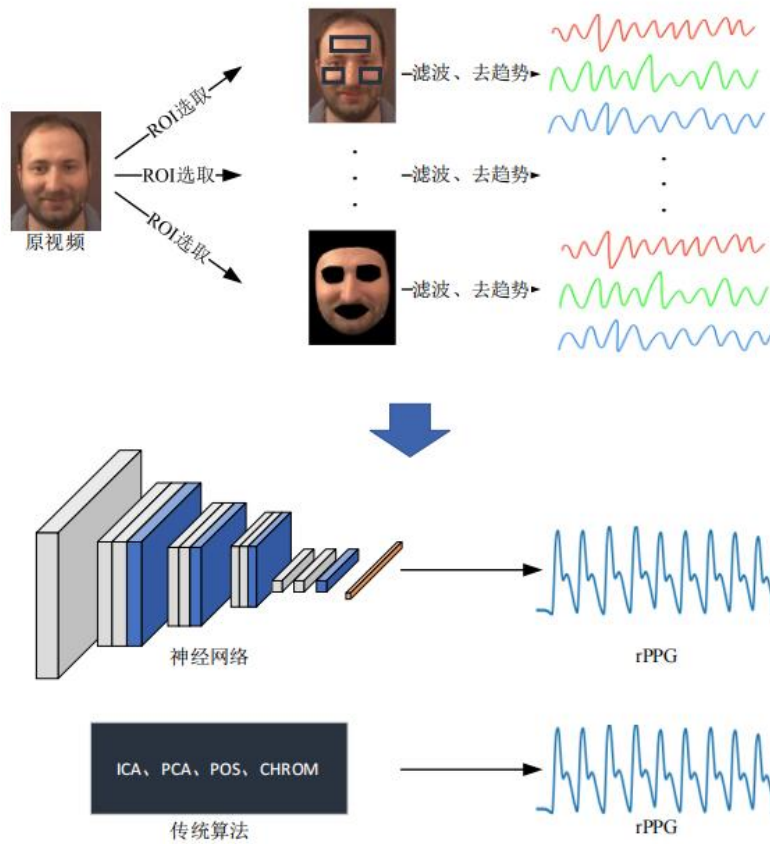


图 2-4 非接触式心率监测流程图

对于基于视觉的非接触式心率监测，面部感兴趣区域的获取至关重要，从良好的感兴趣区域中获得质量较好的周期信息，决定着心率监测数值的准确性。然而，当面对目标移动这种非理想状态下的场景，面部区域往往难以获得具有鲁棒性的信息，同时，摄像头与目标的距离也影响着监测结果，这些都是基于视觉的非接触式心率监测算法无法很好解决的。

2.4 移动目标心率实时监测算法框架

为了保证算法在监测心率过程中采集到的图像清晰可辨以及能够完整的采集运动目标的感兴趣区域，方便后期对采集信息进行分析处理，需要对整个非接触式心率监测算法框架进行改进，保证心率监测算法的有效性和准确性。

基于视觉的非接触式心率监测框架在实际应用中，并不能保证皮肤的感兴趣区域占据大量的像素。导致这个现象的原因可能是主体与成像器距离较远，成像器的像素密度较低，或者视频已经在空间上减少（帧内压缩）所需的存储量。在真实场景下，目标移动，与成像器距离未知，获得的视频中感兴趣区域包含较少有利信息，这不利于心率的监测。现有算法无法很好的解决这个问题。如图 2-5 所示，图(a)，(b)，(c)分别为目标距离成像器 0.5m，1m，1.5m 标注的人脸关键点，可以看出随着目标与成像器距离的增加，人脸关键点越来越难定位。

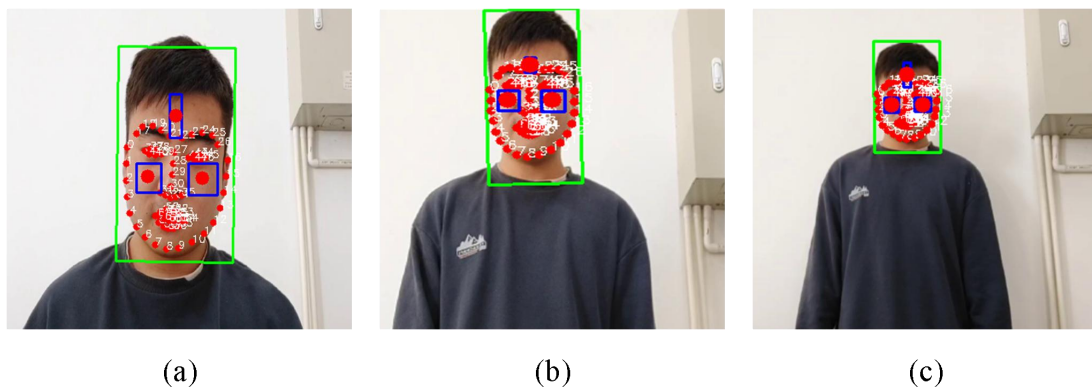


图 2-5 人脸关键点示意图

处于移动状态下的目标，光照路线会发生改变，得到的脉搏波信号中存在一定的噪声信号，被噪声干扰的脉搏波信号不利于心率估计，得到的心率值与真实值有较大出入。运动幅度与频谱中的噪声信号强度成正比。如图 2-6 所示，左图为受到运动干扰的脉搏波信号，右图为没有运动干扰噪声的脉搏波信号，可以看出，受到运动干扰的心率信号的波形更不具有周期性，进而以此得到的心率值误差较大。

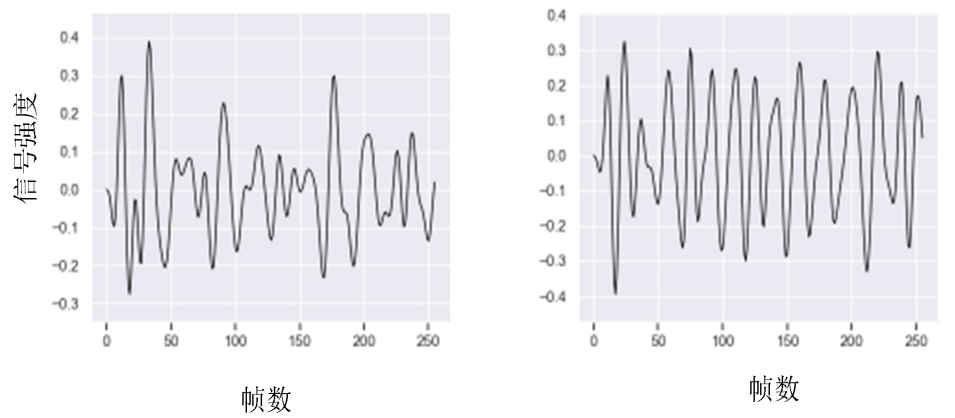


图 2-6 运动干扰脉搏波的信号对比图

本文设计的移动目标心率实时监测框架如图 2-7 所示。

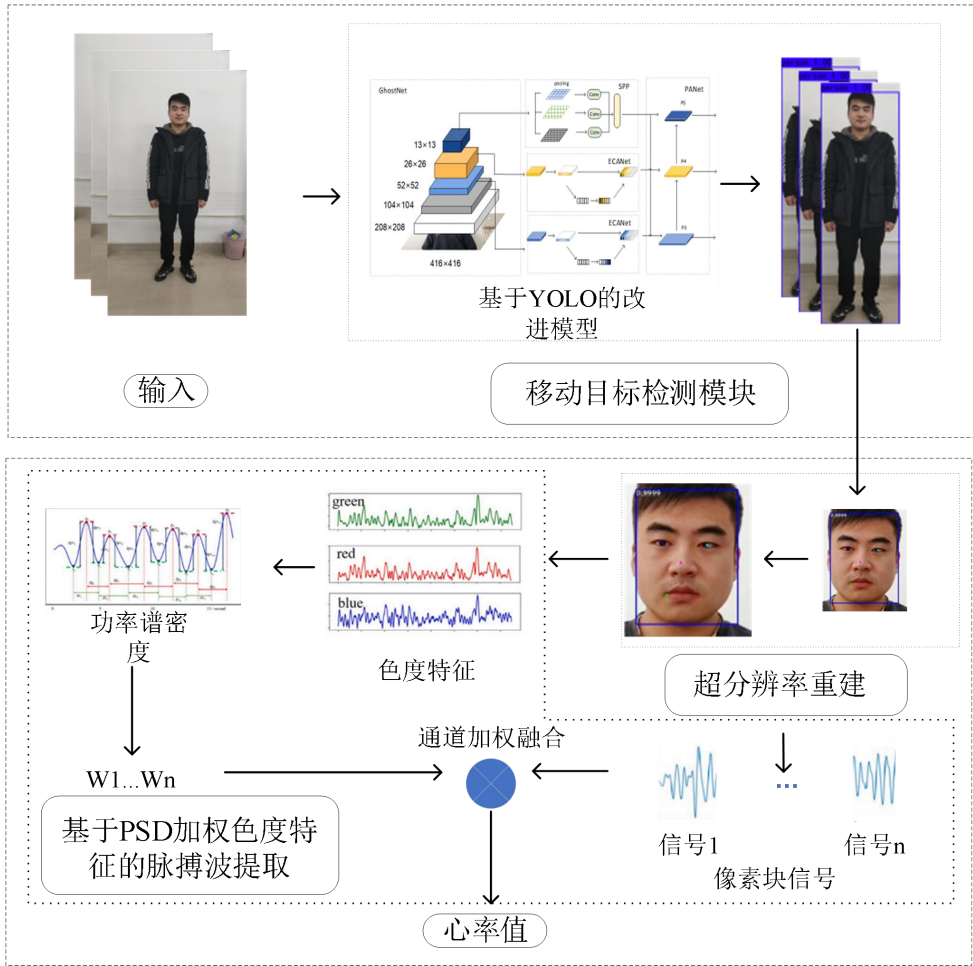


图 2-7 移动目标心率实时监测框架

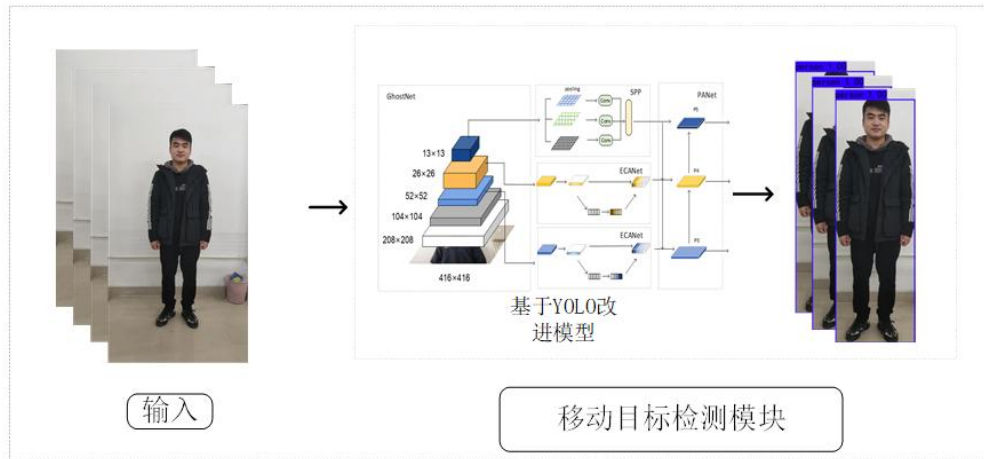


图 2-8 移动目标检测模块

图 2-8 为框架的第一模块，首先会将采集到的图片视频送入该模块，即移动目标检测模块，进行感兴趣区域的获取。当目标距离成像器较远并且处于移动状态时，使用一般方法将无法准确获取感兴趣区域，并且移动目标的运动将会干扰后续心率值信号的提取。在该模块中，选择 YOLOv4 模型为基础进行改进，检测器识别并提取移动中的目标，然后经过处理将所选取的目标框分割，对于每一帧视频都进行此操作，得到一个紧跟被测目标的视频，缩小感兴趣区域，有利于后续心率值的估计。而选择的原始模型在目标移动场景下使用，仍存在问题，例如参数量较大，运行速度较慢，检测精度较低等。针对该模型我们使用轻量化的 GhostNet 替换原本的骨干网络，大大降低了模型的参数量。并且引入高校通道注意力机制，在复杂的输入信息中聚焦于局部信息，将关注更多地放在有利于检测的信息上，忽视没有用的信息，以此降低操作复杂度，提高整个模型的运算准确性和运算速度。

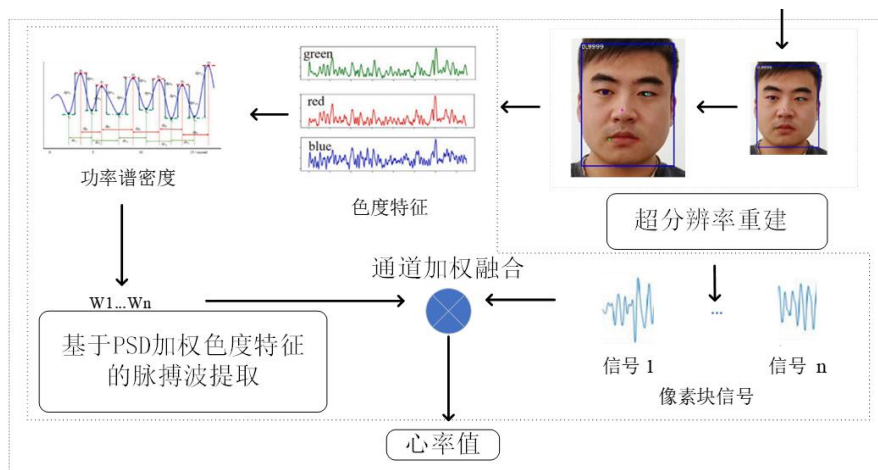


图 2-9 心率检测模块

第二模块是心率检测模块，如图 2-9 所示。首先将会对输入的视频进行预处理。鉴于距离成像器较远，感兴趣区域缩小，此时空间分辨率较低不利于后续心率监测，并且由于设备的影响，得到的视频可能经过压缩处理，这些都会导致视频丢失的相关生理信息。所以在心率检测流程起始阶段，运用超分辨率技术，恢复部分丢失的生理信息，从而增多感兴趣区域内包含的可利用特征信息。接着进入心率监测算法。考虑到人体运动产生运动伪影，将会干扰我们得到准确的 rPPG 信号，在进行原始脉搏波的提取处理步骤时，选择对运动更有鲁棒性的色度特征，并引入 PSD 频域权重，对面部各区域加权融合，最终得到精确的心率值。

2.5 本章小结

本章主要完成了对移动目标心率实时监测框架的研究，对生理学基础进行分析，指出脉搏波的形成和波形特点，其次分析血液容量变化对于光照吸收的光学原理，剖析技术的生物光学原理，总结技术的基本框架。分析了基于视觉的心率检测的框架流程以及对心率监测结果产生影响的干扰因素。介绍了本文提出的移动目标心率实时监测的框架以及设计的原理，依据和优势，为后续模块介绍奠定基础依据。

第3章 移动目标检测方法研究

3.1 本章引言

上一章我们对 rPPG 技术的基本原理、技术框架和影响心率估计的干扰因素进行研究探讨。基于上述基础，提出了本文所运用的非接触式心率监测框架，该框架通过优化关键环节和添加辅助模块来改善算法性能。本章将介绍移动目标心率监测算法的第一模块内容，通过目标检测首先对目标的感兴趣区域进行检测并提取移动中的目标，缩小感兴趣区域，提高监测精度，解决因为与成像器距离较远导致的无法进行心率监测的问题。本模块选择使用基于神经网络的目标检测方法。虽然基于神经网络的目标检测发展迅速，但是应用于本课题的场景中还有不足之处需要改进。如：（1）目前可利用的检测模型参数量较大，运算较复杂，很难在端侧设备上部署，整个算法模型实时性较差。（2）远距离检测人体目标导致小尺度目标关注度变低，存在遮挡情况进而出现目标特征语义信息不足，检测精度降低。为了提升小尺度目标检测精度的同时不会大幅度增加模型参数量，本章提出了基于 YOLOv4 的轻量化移动目标检测算法，即 Little-YOLOv4。

3.2 YOLOv4 算法原理

YOLOv4 模型由三个部分组成，如图 3-1 所示。YOLOv4 模型第一部分是 CSPDarkNet53，其中包括 CBM 模块和 5 个 CSP 模块。首先会将数据集信息放入第一部分，通过该部分将会提取出三个尺度的特征，然后将它们放入空间金字塔池(SPP)和路径聚合网络(PANet)进行特征融合，以此来增强特征信息。YOLOv4 将 SPP 网络、FPN 和 PANet 相结合。SPP 网络的核心作用是增强模型的感受野范围，而 FPN 则运用了自上而下的机制，有效地传递高层次、强语义的特征信息。相反，PANet 采取自下而上的策略，着重传输强有力的定位特征。这三种结构相互融合，能够从多种主干网络层级中汇集特征，并将这些特征逐级扩散至输出层的三个关键特征部分，从而实现对大、中、小目标检测性能的全面提升。

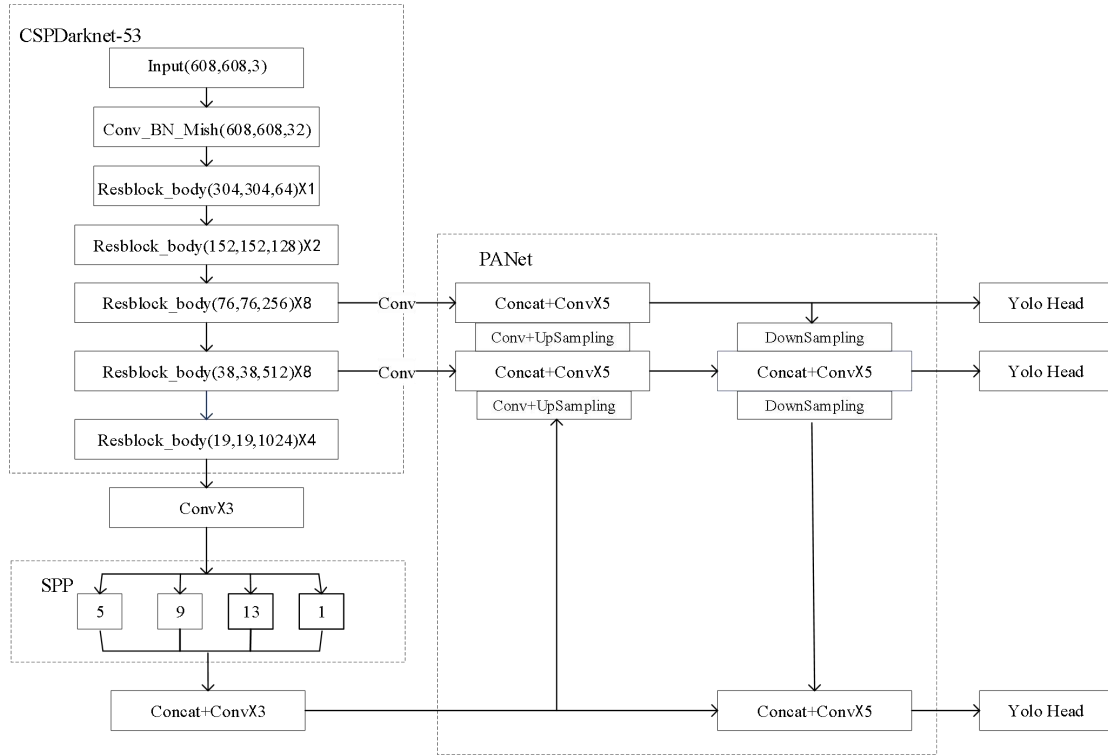


图 3-1 YOLOv4 模型

在预测阶段，YOLO Head 模块构成了预测输出组件，它负责处理经由特征抽取网络和特征融合网络生成的不同尺度特征图，以分别针对小型、中型和大型目标执行检测任务。在具体的预测过程中，对于边界框的预测损失函数，首先采用 Lciou 选取各类得分超过预设阈值的框及其对应得分，继而利用 Ldiou 进行非极大值抑制操作，以筛选出最终的预测框结果。

3.2.1 特征提取网络

特征提取是图像处理中的关键步骤，其核心在于实施数据维度压缩，将原始图像信息转换为具有辨别力且低维的特征向量形式。这一过程始于对图像初级特征的捕获，如早期网络层通常关注图像的基础纹理特性和轮廓线条；随着网络深度的增加，则过渡到提取更复杂、更具抽象意义的高层特征。在 YOLOv4 目标检测架构中，所采用的关键特征抽取子系统是 CSPDarknet53，该网络致力于整合并提炼不同层次的图像特征，以便于模型精准高效地识别各种尺寸的目标物体。

CSPDarkNet53 内部构建包含残差结构单元，其内核设计采用了 CSPNet 理念，如图 3-2 所示，这一设计背后的创新思路在于：原始特征图的一部分通过卷积运算后，利用交叉 stage 连接模式与另一部分特征进行深度融合。这样的构造策略有力地削减了计算负载，压缩了内存占用，同时也显著增强了整个网络的学习能力。

习效能和效率。

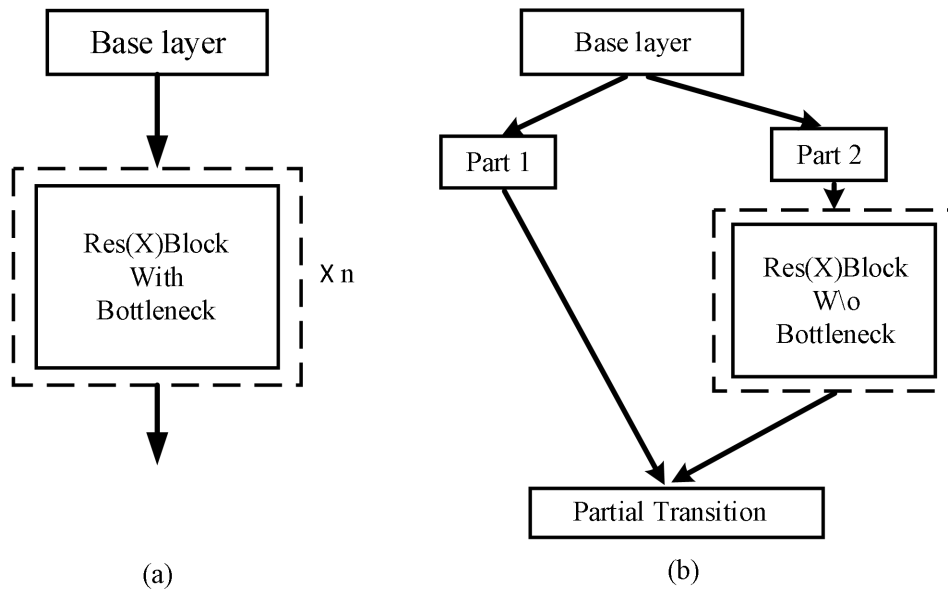


图 3-2 CSP 思想应用到 ResNet 和 ResNeXt 示意图

3.2.2 空间金字塔池化

空间金字塔池化结构(Spatial Pyramid Pooling), 即 SPP, 结构如图 3-3 所示。输入到 SPP 模块的图像尺寸不统一, 不利于后续检测, SPP 模块首先就是要解决这个问题, 融合不同大小的最大池化层, SPP 算法通过将不同语义的特征融合, 以此来提高模型的适用范围, 可以检测更多不同大小的目标。将该模块加入算法, 是为了扩大模型可以计算得到的感受野范围。YOLOv4 中的 $k = \{1*1, 5*5, 13*13\}$ 。当 SPP 模块被集成到物体检测算法中时, 其重要作用体现在仅需对输入图像进行一次特征提取映射, 即可将其有效压缩并转化为一个固定输出的子图像表示。这样一来, 便巧妙解决了传统方法中卷积特征可能需要多次重复计算的问题, 从而提高了算法运行效率。与使用 $k*k$ 最大池化的方式相比, SPP 模块只需要计算一次卷积层, 训练速度快, 更有效的扩大了模型接收的特征尺寸, 有效的加强了特征提取能力, 分离出了利于后续检测需要的上下文特征。

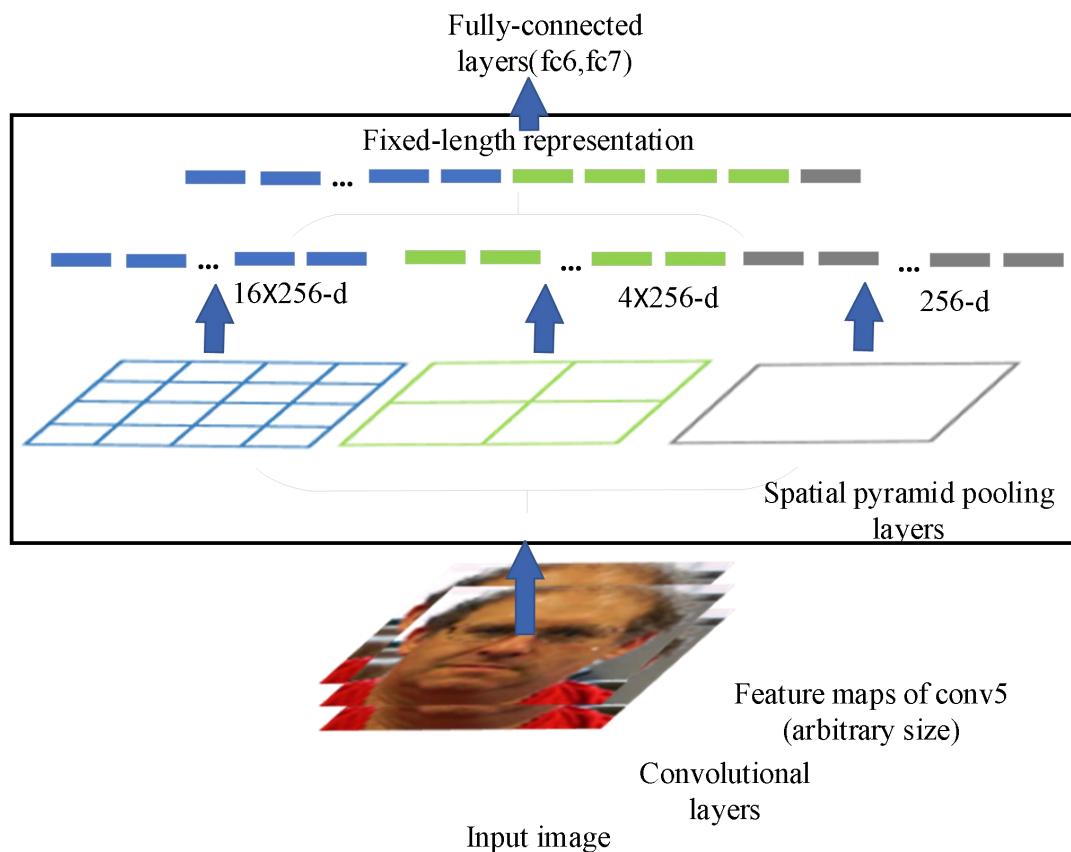


图 3-3 SPP 模块原理图

3.2.3 路径聚合网络

路径聚合网络的结构如图 3-4 所示，其中(a)是 FPN 结构(b)是 PAN 增加的自底向上的特征融合层。其中(a)部分借鉴了图像金字塔的主要思想，将通过主干提取网络得到的特征图进行处理，将高层语义信号向低层融合。就可以构建得到金字塔特征图，将提取得到特征，自上向下，将高层的特征传下来。(b)引入自底向上的捷径，使得底层信息更方便容易的传到顶层。(c)为自适应特征池化操作，聚

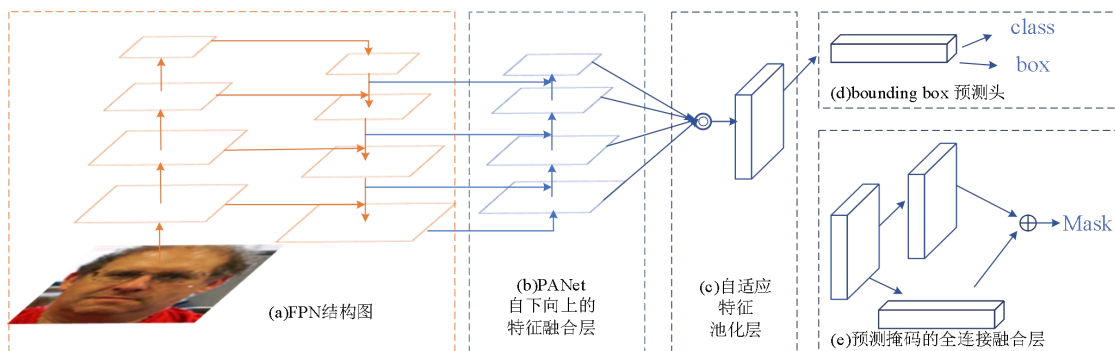


图 3-4 路径聚合网络的结构

合了多层的信息，对定位更加有利。(d)进行全连接层融合。路径聚合网络的结构特殊的改进点在于 PAN(Path Aggregation Network)结构进行特征融合时并不是将两个相同大小的特征映射执行按位加操作，而是进行拼接操作，如图 3-5 所示。(d)是该部分的预测头，(e)是全连接融合层。通过组合这几个模块，可以很好的完成目标定位的功能。

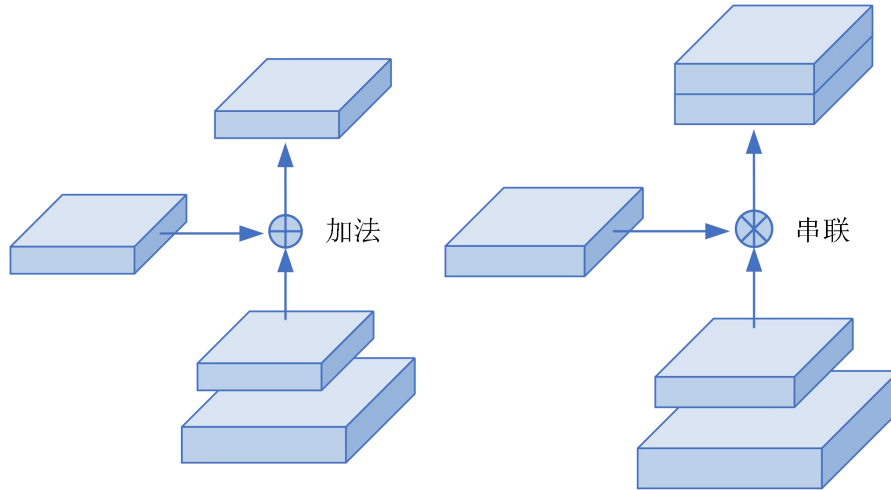


图 3-5 路径聚合网络特征堆叠示意图

3.3 基于 GhostNet 的移动目标检测改进

3.3.1 GhostNet 模型

GhostNet 模型为了去除在得到特征图过程中冗余的部分操作，使用一些计算量更低(Cheap Operations)的操作，使用这个操作将会有效的降低模型的参数量，降低模型复杂度，同时模型的性能并未降低，模型运行速度更快。

从整体上看 Ghost-Module，简单来看就是两步：即：（1）首先压缩输入的特征层的通道数，这里使用 1×1 的卷积进行，获得局部特征信息浓缩。（2）之后利用多个卷积层获得多层特征图，然后将特征层进行堆叠处理，利用深度可分离卷积获得特征浓缩的相似特征图。从输入提取到特征层中的特征信息包括很多无用或重复的，一般的卷积操作会对相似的信息重复操作，这样操作需要耗费大量的流程步骤和参数，这里使用的 Ghost 模块替代了传统的卷积。在 Ghost 模块中，首先通过标准卷积操作提取出信息相对稀疏的特征图层，随后运用简化版的计算过程进一步生成数量更多的特征图谱。如图 3-6，3-7 所示。

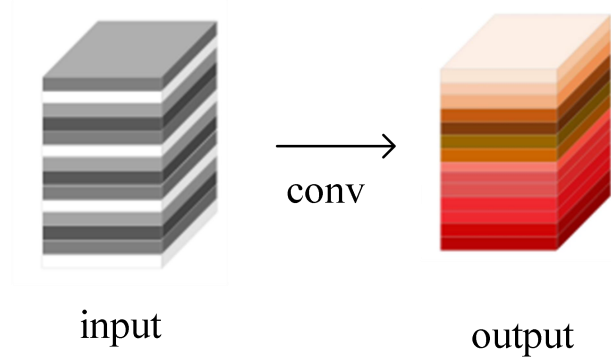


图 3-6 普通卷积示意图

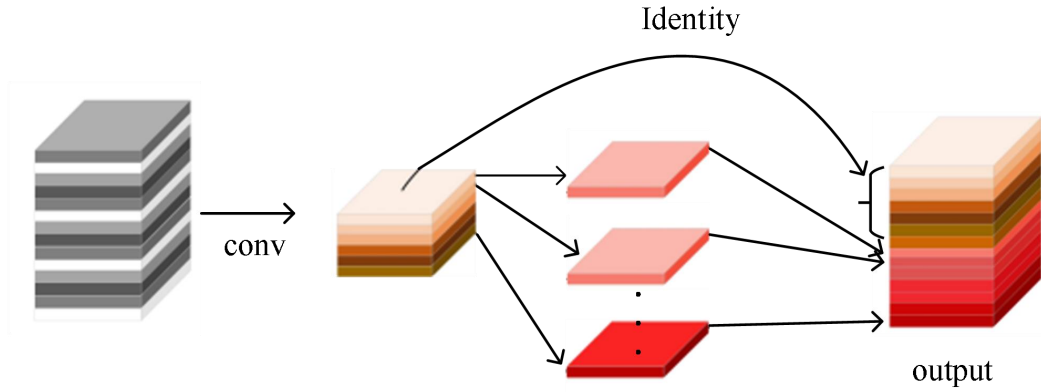


图 3-7 Ghost 模块示意图

Ghost Bottlenecks 是一种由一对 Ghost Modules 构成的独特瓶颈设计。其存在两种变体，如图 3-8 所示：(a)在该结构的第一种形式中，Ghost Bottleneck 嵌入在主干网络内部，其中包含两个连续的 Ghost Module，专注于对输入特征层进行特征抽取。不同于其他部分，其残差连接不进行额外处理，而是直接将原始输入特征与经过 Ghost Module 处理后的输出相加。由于此瓶颈结构的步长设置为 1，故不会对输入特征层的高度和宽度进行任何压缩操作，这种设计有助于增强网络的深度。(b)另一种类型的 Ghost Bottleneck 同样由两部分组成。不过，在这一变体中，主干部分首先应用一个 Ghost Module，随后采用步长为 2 的深度可分离卷积来对输入特征层进行高度和宽度的缩小处理，紧接着再通过另一个 Ghost Module 进一步进行特征提取。由于此处瓶颈结构的步长为 2，因此能对输入特征层进行尺寸缩减，这种结构设计便于调整特征层的形态。

输入特征图的尺寸为 (w, h, c) ，输出的特征图的卷积核为 $k * k$ ，通道数为 n ，线性变换核为 $d * d$ 。普通卷积的计算量为：

$$\text{cost} = h' * w' * n * k * k * c \quad (3-1)$$

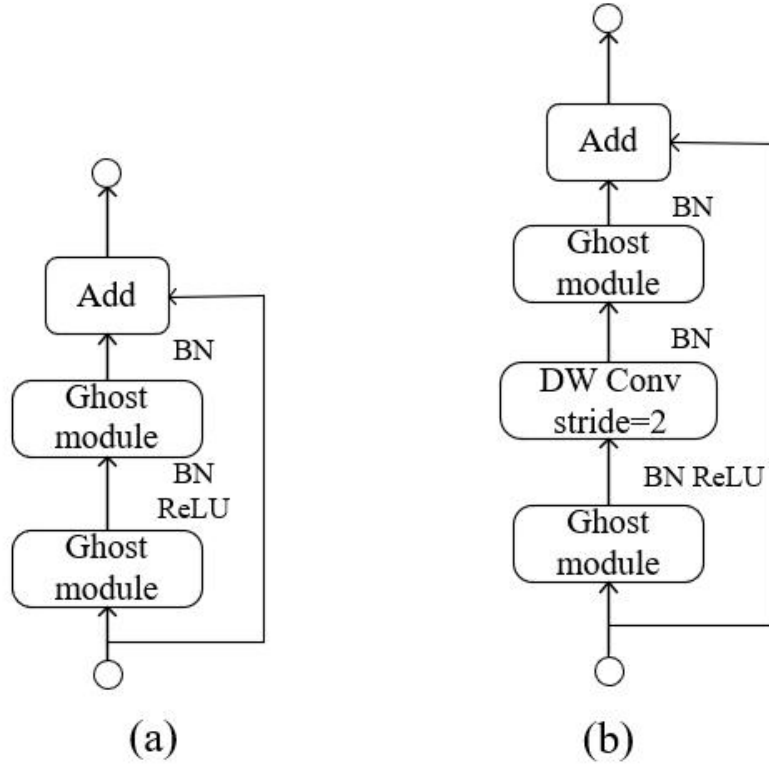


图 3-8 Ghost Bottlenecks 结构

而提出的 Ghost Module，其计算量为：

$$\text{cost} = h' * w' * \frac{n}{s} * k * k * (c+s-1) \quad (3-2)$$

两者之比为：

$$r_c = \frac{n * c}{\frac{n}{s} * (c+s-1)} = \frac{s * c}{c+s-1} \approx s \quad (3-3)$$

由以上推算可以知道，提出的 Ghost Module 将极大的减少模型的参数量。

3.3.2 基于 GhostNet 的 YOLOv4 改进网络

YOLOv4 模型使用特征提取网络模型 CSPDarknet-53，而 CSPDarknet-53 包含较多参数，下采样次数过长，这不利于快速收敛，网络也存在退化的可能性。因此在本文中我们将原始模型的特征提取部分进行替换。选择前文介绍的 GhostNet 替换 YOLOv4 模型中的原始骨干提取部分。降低整个模型运算所需要的参数量，降低了模型的复杂度。

在改进后的 YOLOv4 网络体系中，特征融合环节沿用了经典的 FPN（特征

金字塔网络) 结构, 旨在对由 GhostNet 模型产出的三个独立特征流进行深度整合与交融, 以此实现多层次、多尺度的特征信息的有效利用, 进一步轻量化模型。将原始模型 YOLOv4 中的特征提取网络 CSPDarknet-53 替换成 GhostNet。结构参数如下表 3-1 所示。

表 3-1 YOLOv4+GhostNet 结构

Operation	step size	input	Output	nucleus	step length
Conv2D_ReLU		(416,416,3)	(208,208,16)	(3,3)	2
G-bneck1		(208,208,16)	(208,208,16)	(3,3)	1
G-bneck2		(208,208,16)	(104,104,24)	(3,3)	2
		(104,104,24)	(104,104,24)	(3,3)	1
G-bneck3		(104,104,24)	(52,52,40)	(5,5)	2
		(52,52,40)	(52,52,40)	(5,5)	1
G-bneck4		(52,52,40)	(26,26,80)	(3,3)	2
		(26,26,80)	(26,26,80)	(3,3)	1
		(26,26,80)	(26,26,80)	(3,3)	1
		(26,26,80)	(26,26,112)	(3,3)	1
		(26,26,80)	(26,26,112)	(3,3)	1
G-bneck5		(26,26,112)	(13,13,160)	(5,5)	2
		(13,13,160)	(13,13,160)	(5,5)	1
		(13,13,160)	(13,13,160)	(5,5)	1
		(13,13,160)	(13,13,160)	(5,5)	1
		(13,13,160)	(13,13,160)	(5,5)	1

整个 Ghostnet 都是由 Ghost Bottlenecks 组成的。输入特征图的尺寸为 (416,416,3), 首先进入第一个瓶颈结构, 该结构步长为 1, 通道数与尺寸均不发生变化; 之后进入 G-bneck2, 该部分有两个瓶颈结构, 第一个步长为 2, 进行高和宽的压缩, 输出为(104,104,24), 第二个步长为 1 不进行压缩, G-bneck2 输出尺寸即为(104,104,24); G-bneck3 同样包括一个步长为 2 和一个步长为 1 的瓶颈结构, 特征层输出尺寸为(52,52,40); G-bneck4 首先经过一个步长为 2 的瓶颈结

构，压缩宽和高，再经过 4 个步长为 1 的瓶颈结构处理，输出尺寸为(26,26,112)；G-bneck5 与 G-bneck4 结构相同，输出尺寸为(13,13,160)。经过多个 Ghost Bottlenecks 的堆叠，最终获得了一个(13,13,160)的特征层。

3.4 基于高效通道注意力机制的移动目标检测改进

3.4.1 高效通道注意力机制

高效通道注意力机制模块(Efficient Channel Attention, ECA)是在通道注意力机制(SE)的基础上进行改进提出的。在研究 SE 通道注意力机制模块时，发现模块中的降噪操作不利于学习通道注意力，并且获取依赖关系效率较低。

ECA 模块的整体结构与原理基本与 SE 模块相同，结构图如图 3-9 所示。SE 模块使用两个非线性完全连接(FC)来捕获通道间的非线性相互作用，同时引入降维步骤以减少模型的整体复杂性，而这个降维操作是低效的。ECA 模块的特点在于研究得到通道与其权值是存在一定关联性的，这里找到了两者的关系，即公式 3-5，通过得到输入图像的通道数进而利用公式得到一维卷积核大小 $kernel_size$ ，该步骤称为激励操作。通过这个关系就可以知道通道数对应的权值，最后将权重和通道一一相乘，就可以得到具有权重的特征图。激励操作降低模块的参数量的同时提高机制的学习能力。

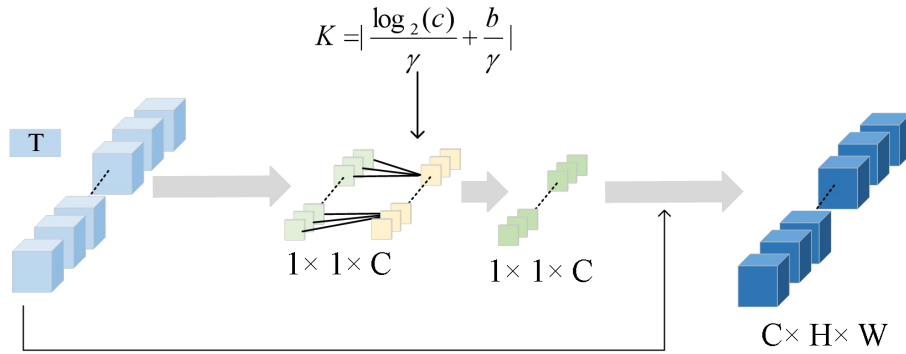


图 3-9 ECA 模块示意图

ECA 模块的激励过程计算公式为：

$$w = \sigma(\text{Conv}_k(y)) \quad (3-4)$$

卷积和自适应函数，定义如下：

$$k = \psi(C) = \left\lceil \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rceil_{odd} \quad (3-5)$$

其中 k 为卷积核大小， C 为通道数， $\lceil \cdot \rceil_{odd}$ 为最近取到的奇数，算法中的 k 值是可以手动调整的， k 值的大小关系到可以提取有效跨通道信息的数量， γ 和 b 都是设定好的常量。

3.4.2 基于高效通道注意力机制的 YOLOv4 改进网络

YOLOv4 模型的特征融合部分采用 FPN 搭建的阶梯式的网络。在特征提取网络内部，设计了并行连接路径，该结构能够直接实现高层特征与底层特征的同时提取与融合操作，有助于融合高层特征与底层特征。原始的卷积神经网络虽然可以获得图像的特征信息，但无法对于一些重点信息的获取能力还不够，这样不利于检测器的检测。所以在原始模型中引入高效通道注意力机制(Efficient Channel Attention, ECA)，提高模型检测精度，同时提高检测速度。

传统的注意力机制有 SENet, ECANet 等。其中，SENet 主要是包括 squeeze 和 excitation 两大操作，其中降维处理在实现通道注意预测的过程中带来了不利影响，获取所有通道之间的关系是没有必要的，也是降低整个模型效率的操作。在前文已经将两者进行比较，ECANet 使用自适应的一维卷积实现跨通道的交互，而 SENet 则是采用全连接和激活函数处理特征向量，使用 ECANet 将会使得模型参数量大幅减少，提高模型的效率。

因此本课题将在原始模型中引入 ECA 注意力机制。注意力机制集成的位置关乎整个模型运算的复杂度以及模型的精确度，它可以放在任何特征图后面，但是将放在特征提取部分，会破坏整个模型的结构，并且模型经过多轮训练得到的权值也不可利用，这样将会造成整个操作变得低效。因此改进后的模型在非主干部分加入了 ECA 模块。ECA 注意力机制以及 GhostNet 集成到 YOLO 网络中的结构如图 3-10 所示。

心率监测框架的第一模块，基于 YOLOv4 进行改进。改进后的网络结构图如图 3-10 所示，下面将详细介绍该模块的流程步骤以及改进后的网络结构。首先预处理阶段使用 Mosaic 数据增强，该方法是在 CutMix 基础上进行改进得到的。Mosaic 数据增强利用随机的四张图片，将它们做一些简单的处理，即旋转，缩

放等，然后将处理后的图片拼接，就会得到一个全新的图片，将其送入网络学习，这样操作相当于一次送入四张图片进行学习，提高了学习的效率。Mosaic 数据增强不直接增加 `batch_size` 的同时，极大的丰富了检测目标的背景，有助

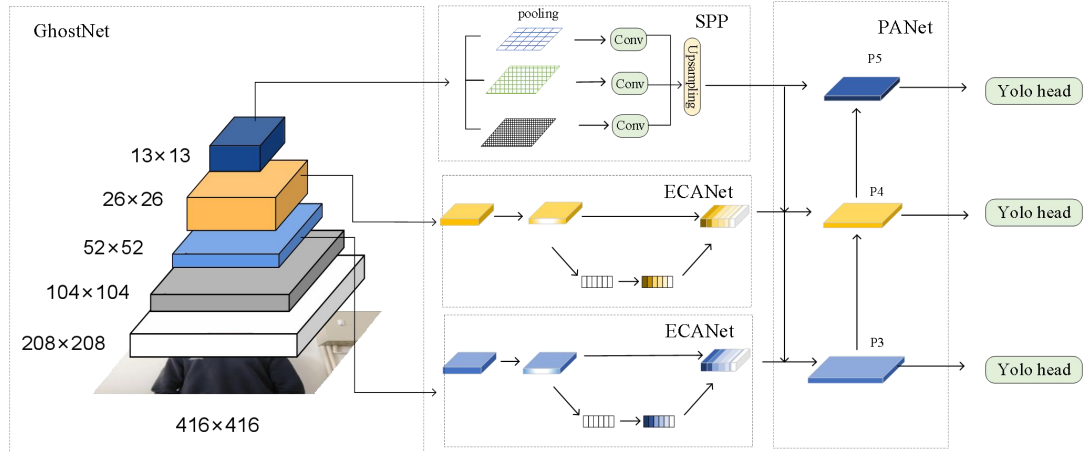


图 3-10 改进后的网络结构图

于后续准确性的提升。然后进入特征提取步骤，该部分将 YOLOv4 主干网络部分换成 GhostNet 模型，用于初步的特征提取。颈部检测网络由 SPP 和改进的 PANet 组成，用于进一步的数据特征提取；同时，引入 ECANet，嵌入到 YOLOv4 检测网络中，大大提高了提取有效信息的能力，进一步减少了计算量。

3.5 实验过程与结果分析

3.5.1 数据集和实验环境

移动目标检测模块实验过程，是用 Python 在 NVIDIA RTX3090GPU 上使用 Pytorch 实现的。使用 VOC2007 数据集进行训练测试，此数据集能够进行深度学习的训练，主要研究目标为移动的行人，因此只保留 Person 类图片，共 9963 张图片。实验使用图像数据集如下图 3-11 所示。

实验中，设置 Batch size（一次批量训练的数据样本数量）为 16，Image size（图像大小）为 416×416 ，Epochs（整个数据集的训练次数）为 200，选择的优化器为 `sgd`，模型的最大学习率为 $1e-2$ 。实验所用到的硬件配置在表 3-2 中说明。

表 3-2 实验所用硬件配置

设备	配置
系统	Windows10
CPU	Intel Core i3-12100F 3.30 GHz
GPU	NVIDIA GeForce RTX 2060



图 3-11 VOC2007 数据集

3.5.2 评价指标

在移动目标检测领域中,评价模型效能的常用指标是准确率和召回率,而这两种性能度量的计算均离不开混淆矩阵的支持。下面为分类器的准确率(Precision Rate)和召回率(Recall Rate)计算公式:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3-6)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3-7)$$

$$AP = \int_0^1 P(R) dR \quad (3-8)$$

式中: TP 表示样本被分为正样本且分配正确,FP 表示样本为负样本且分配正

确，FN 表示样本被分为负样本但分配错误。

3.5.3 实验结果分析

将本文提出的算法与多个传统的算法进行性能比较，使用经过筛选的 VOC2007 数据集。结果如下表 3-3 所示，其中 YOLOv3 的平均精度只有 73.81%，表现较为不错的 SSD 模型的平均精度为 81.54%，仍低于文本模型。而 YOLOv4-Tiny 是根据 YOLOv4 改进的轻量化模型，表现较差。

表 3-3 VOC 数据集在各类别检测器 AP 值

模型	AP(%)
YOLOv3	73.81
SSD	81.54
YOLOv4-Tiny	68.70
RetinaNet	81.70
Little-YOLOv4	82.91

为验证每个模块的作用，设计消融实验进行验证。图 3-12 为实验结果，(a) YOLOv4 (b) YOLOv4+GhostNet (c) YOLOv4+GhostNet+EcaNet。GhostNet 轻量级模型，在于减少模型参数，所以精度相比 YOLOv4 本身有所下降。由实验结果可知，使用了 ECANet 注意力机制后，比起原始模型，精度提升 0.05%，而参数量大幅度降低。如表 3-4 所示，比较了本文算法和原始检测器的总体参数。可以看出，YOLOv4 的参数为 64363101，而本文的总参数约为 11428679，占 YOLOv4 的 17.7%，明显小于原始检测器。这可以证明我们提出的轻量级方法可以在保持精度的同时显著减少原始模型的参数总数。

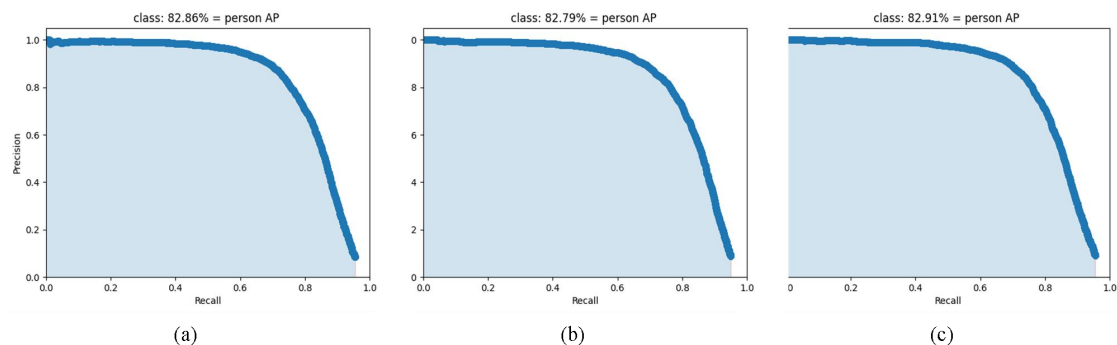


图 3-12 PR 曲线图对比

表 3-4 模型参数量

模型	参数量
YOLOv4	64363101
YOLOv4+GhostNet	11428545
Little-YOLOv4	11428679

YOLOv4 原始模型和本文改进后得到的 Little-YOLOv4 模型的检测效果图如下所示:



图 3-13 YOLOv4 检测结果示例图



图 3-14 Little-YOLOv4 检测结果示例图

3.6 本章小结

本章介绍移动目标心率监测算法的第一模块内容。首先使用目标检测模型定位感兴趣区域并提取移动中的目标。原始模型应用于本文存在一定局限性，于是基于原始 YOLOv4 网络进行改进，引入 GhostNet 和 ECANet。最后使用经过筛选的 VOC2007 数据集，与其他模型进行对比，在一定程度上提升了模型的性能，为之后的心率监测奠定一定的基础。

第 4 章 基于 PSD 加权色度特征的心率估计方法研究

4.1 本章引言

上一章我们对移动目标检测方法的基本原理,影响因素进行研究探讨,得到了改进后的移动目标检测方法。以 YOLOv4 模型为基础,通过优化各环节来改善性能,选用了多个指标来评价效果。在心率监测模块,本章节着重探讨了基于视觉的监测算法应用到移动目标场景时遇到的挑战,聚焦于算法优化工作,通过对关键步骤改良,削弱外部环境因素对生理参数提取所造成的负面影响,从而提高整个算法的综合性能。

目标处于移动状态,给感兴趣区域的定位以及关键信息的获取带来挑战。针对这些问题,本章提出了一种基于 PSD 加权色度特征的心率估计方法。首先进行视频预处理步骤,使用超分辨率网络恢复丢失的生理信号,其次选择色度特征进行二次去噪处理,引入 PSD 计算面部各区域的对应权重,最后进行信号加权。通过这些操作一定程度上的去除现实环境中目标运动的干扰,之后将会得到较为纯净的心率信号,最终得到心率估计值。在最后一节将会通过实验,验证该模块的有效性,在公开数据集上与其他传统算法进行性能比较。

4.2 基于输入优化的 rPPG 心率检测

目前,各种视频压缩方式不断被提出,这些技术虽然可以一定程度上的减少视频数据量,但是这些操作极大的降低了视频的比特率。现有的这些技术大多都是假定同一图像上,各像素点的色彩差异很小,图像之间的差异也很小,之后通过舍弃部分细微信息的方式,达到压缩视频所含数据量的目的,但此类操作不可避免地将会丢失视频中部分生理信号。

并且,在本文课题场景下成像器帧率和分辨率、被测者到成像器的距离、视频压缩、感兴趣区域缩小等变量都可能会影响通过 rPPG 方法得到的任何生理测量的准确性。移动目标在视频中发生位移,通过框架的第一模块处理,我们将大部分移动影响排除掉了,同时又带来一个新的问题,视频的质量将会变差。低分

分辨率的视频质量较差, 视频中可提取的信息少, 难以得到心率估计所需要的信号, 并且这些特征的提取极易受到外界因素的干扰, 给后续的心率检测阶段带来不利影响, 难以获得精准的测量心率。为此, 本节首先进行面部视频的信号优化工作。本文将在心率监测模块添加预处理步骤, 恢复视频中被丢失掉的生理信号。

4.2.1 基于传统超分辨率重建模型的输入优化

下面介绍并分析了几种传统的超分辨率重建模型, 将会解释它们的原理思想以及对应存在的优缺点。给后续预处理步骤的模型选择奠定理论基础。

最邻插值法, 利用了图像处理中的插值函数, 通过对原始低分辨率图像插值, 获得高分辨率图像中的对应像素值。而对于双线性插值法来说, 该方法依赖于一个涉及两个变量的插值函数, 其基本原理是在二维空间中分别进行两次一维线性插值, 即先在一个轴向上进行两次线性插补, 然后在另一个轴向上进行一次线性插补, 从而实现像素值的精确估计。总共做三次插值得到最终解。如图 4-1 所示。

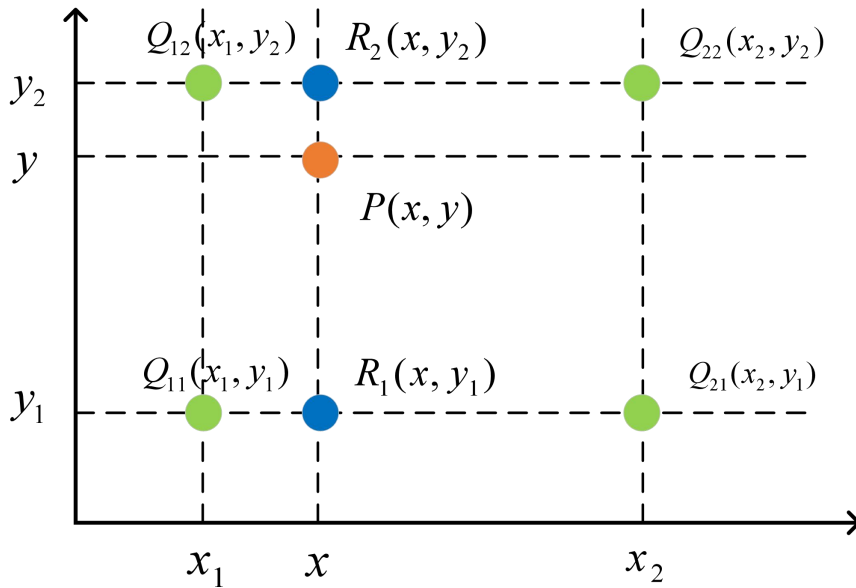


图 4-1 基于插值的超分辨率方法

POCS 算法, 基本思想是在向量空间中的任意起始点出发, 搜寻并确定一组满足高分辨率图像重建约束的所有可能解, 这些解位于预设的约束集合之内。尽管此类方法概念简洁、计算负担较小, 但在面对结构复杂的图像问题时却显得力不从心。另外, 在操作流程中采用的插值手段易造成高频信息的流失, 由此产生的副作用就是使重构后的图像边缘呈现模糊不清的状态。

4.2.2 基于 Real-ESRGAN 模型的输入优化

由于本文选择的是基于面部视频的心率监测算法，需要用到的大量 rPPG 信号均来自面部，所以我们将研究的重点放在优化输入图像上。通过这样操作，可以降低模型所需的拟合数据量，从而降低了模型的负荷，使得后续的监测心率部分效率较高。本文将在心率监测进行前引入人脸图像优化模型，提高面部信息的质量，恢复丢失的部分生理信息，为后续心率的准确测量打下基础。

基于深度学习的超分辨技术相比起传统的方法，它考虑了图像的内在结构和先验信息，能够更好地重建高分辨率图像，而且具有较强的自适应能力，能够根据不同的低分辨率图像自动调整模型参数。它提高了在主观视觉上的效果并且在客观质量评估方面获得不错的成果。

(1) SRCNN

超分辨卷积网络(SRCNN)是卷积神经网络应用在单幅图像超分辨率领域内具有里程碑意义的模型。比起传统的方法，该方法提高了模型精度。SRCNN 模型结构由生成器、判别器和 vgg 网络构成。模型结构如图 4-2 所示。

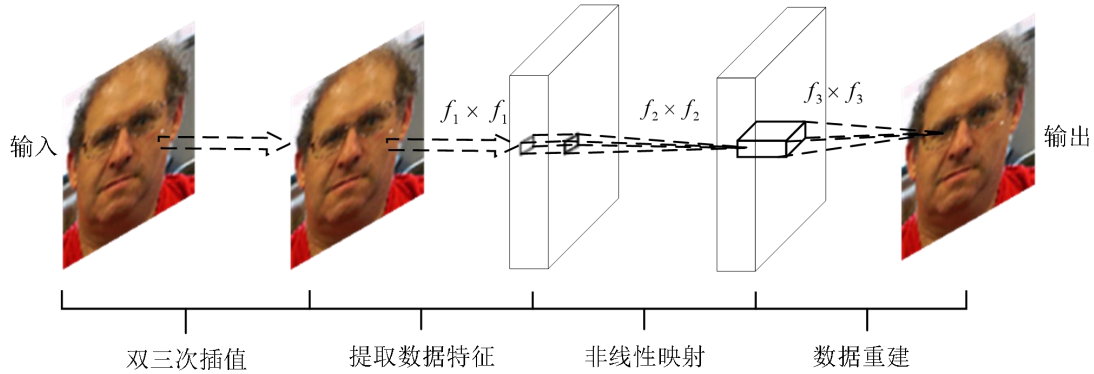


图 4-2 SRCNN 结构图

模型整体由三层卷积层组成。对于输入的低分辨率图像，该模型首先利用双三次插值，该操作是为了将原始图像放大，但是此时由于仅仅是放大，图片仍是低分辨率图像；其次通过三层卷积网络层拟合非线性操作，第一层卷积的作用是将图像中的各个图像块的特征提取并表示出来。卷积核 W 尺寸为 $C \times f_1 \times f_1 \times n_1$ ， c 表示图像的通道数， $f_1 \times f_1$ 表示卷积核的尺寸， n_1 表示该卷积核的维度， B 表示偏置，经过第一层卷积就可以得到 n_1 个特征图。第一层卷积输出就可以表示为：

$$F_1(Y) = \max(0, W_1 * Y + B_1) \quad (4-1)$$

第二层卷积层的目的是得到特征非线性映射。尺寸为 $n_1 \times f_1 \times f_1 \times n_2$ ，输出得到 n_2

个特征图，第二层输出表示：

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2) \quad (4-2)$$

最后一层卷积层使用乘加结合的方式重新构建得到高分辨率图像。卷积核尺寸为 $n_2 \times f_1 \times f_1 \times C$ ，得到输出图像，第二层输出表示：

$$F_3(Y) = W_3 * F_2(Y) + B_3 \quad (4-3)$$

最后，模型最小化损失函数如下式所示：

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \theta) - X_i\|^2 \quad (4-4)$$

其中， F 代表输出， θ 表示优化的参数， X_i ， Y_i 代表给出的高分辨率图像与低分辨率图像。然而，SRCNN 算法首先使用了双三次插值增加了运算量，其次使用三层卷积层，层数较少将会导致模型精度不够，模型仍存在一定缺陷。

(2) Real-ESRGAN 结构

在本文中，选择基于 SRCNN 模型进行几次改进的 Real-ESRGAN 对人脸视频进行处理，用以恢复原始视频的信息损失，以便后续心率监测。

使用该模型是为了恢复被严重腐蚀或者丢失的 rPPG 信号。模型选择各项性能较为均衡，且使用纯粹合成数据集训练，展现出强大的能力，能够有效地复原真实世界中的图像的 Real-ESRGAN 模型。它比起其他传统模型的架构简单，参数量适中，且生成高分辨率图像的性能较好，在实际应用场景中更具有实用性。下面将详细介绍该模型的结构。

Real-ESRGAN 模型是一个生成对抗网络，由生成器 G 和鉴别器 D 组成。生成器是用真实数据训练一定步数，尽最大可能得到逼真的样本。鉴别器 D 的目的是区分出生成器生成得到的图片数据是否为真，判断并找出生成的虚假数据。两者同时训练，两者的能力也不断提高，直到达到平衡。最终会得到的平衡是：生成器 G 生成的数据经过鉴别器 D 判断是真实的，鉴别器 D 也无法区分数据是使用生成器 G 生成的虚假数据还是真实数据。此时 $D(G(z)) = 0.5$ 。其损失函数定义如下：

$$L_{GAN}(G, D) = E_{X \times P_{data}(x)} [\log D_{\theta}(x)] + E_{Z \times P_z(z)} [\log(1 - D_{\theta}(G_{\omega}(z)))] \quad (4-5)$$

$$\arg \min_G \max_D L_{GAN}(G, D) \quad (4-6)$$

上式中 x 代表真实数据值， z 是随机噪声， θ 与 ω 是参数。

本章使用 Real-ESRGAN 作为人脸图像的优化模型。其中，生成器结构如图 4-3 所示，将 RRDB 作为基本块，这里使用 23 个基本块。RRDB 结构如图 4-4，该结构极大的提升了模型特征提取能力，它结合了多级残差网络和密集连接，每个 RDDB 块由 3 个 Residual Dense Block 组成，输入和输出形状一样，然后进行两次 Unsample，Unsample 采用 nearest 插值，每次 Unsample 之后会有卷积层来细化插值细节，最后通道数变成 3 输出。

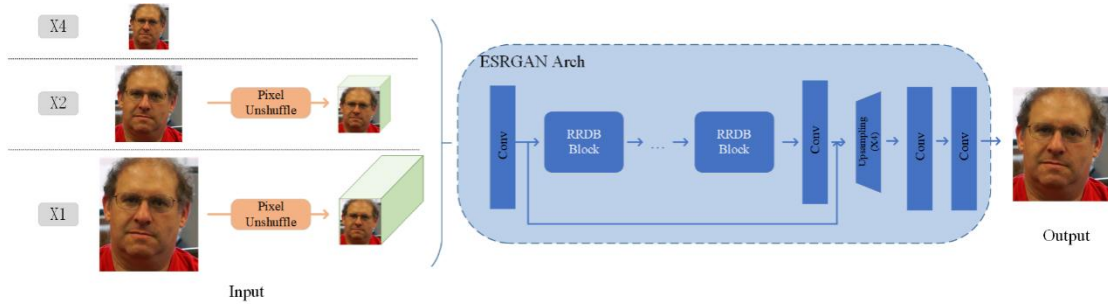


图 4-3 生成器基础结构

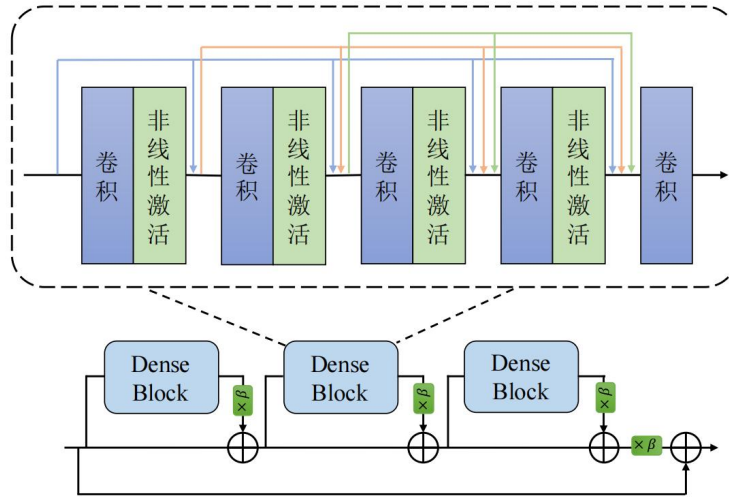


图 4-4 RRDB 结构

Real-ESRGAN 需要使用对复杂输入有更强能力的判别器，这个判别器可对局部纹理有更精确的梯度。所以选择了使用带有谱归一化的 U-Net，它可以对图像的每一个像素进行判断，结构如图 4-5 所示。

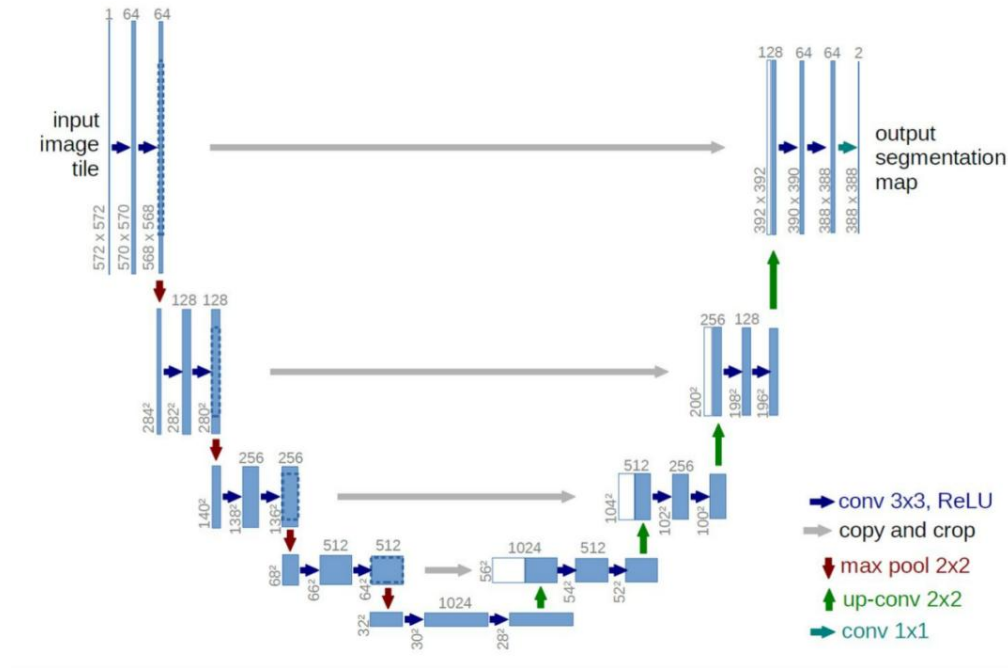


图 4-5 U-Net 结构图

如上图, U-Net 网络结构是由一个 Encoder 和一个 Decoder 组成的对称结构。Encoder 是图中左部分, 包括模型的特征提取以及下采样部分, Decoder 是图中右部分, 该部分将通过一系列上采样操作进行解码。以输入一张单通道图像为例, 首先使用两个卷积核为 3×3 的卷积层, 该卷积层步数为 1, padding 为 0, 所以每次卷积之后 feature map 的 H 和 W 都会变小, 然后经过下采样操作 pool 为 2×2 , 图像高和宽减半, 再进行两个卷积层, 之后经过三个同样的操作; 接下来进行转置卷积操作, 与经过中心裁剪的部分拼接, 紧接着进行卷积核为 2×2 的上采样, 然后经过两个 3×3 的卷积层再与经过中心裁剪的部分进行拼接, 之后进行两个一样的操作, 最后通过 1×1 卷积层卷积。U-Net 使用拼接的方式连接, 特征图的维度没有发生改变, 这种方式相比起普通模型, 具有更加高效便捷的优点。

4.3 基于 PSD 加权色度特征的脉搏波信号处理方法研究

4.3.1 Viola-Jones 算法

由于本课题是基于人脸视频提取面部周期性变化形成的信号进行心率检测, 面部的准确提取在一定程度上会影响得到心率值的准确性, 所以在进行心率检测之前, 选择合适的人脸检测方法是采集高质量生理信号的关键。

Viola-Jones 算法较为经典的人脸检测方法。该算法首先使用 Haar-like 特征

表示出目标的面部特征，得到多个矩形特征值，引入积分图像，这样图像的矩形特征将只受到与其对应的积分图影响。这个操作将会加快模型的运算速度，有利于 Adaboost 算法进行筛选。然而最初的分类器不足以检测到准确的结果，于是通过 AdaBoost 算法得到一个强分类器，此时检测结果较为准确。流程如下图 4-6 所示，下面将详细介绍该方法的具体实施细节。

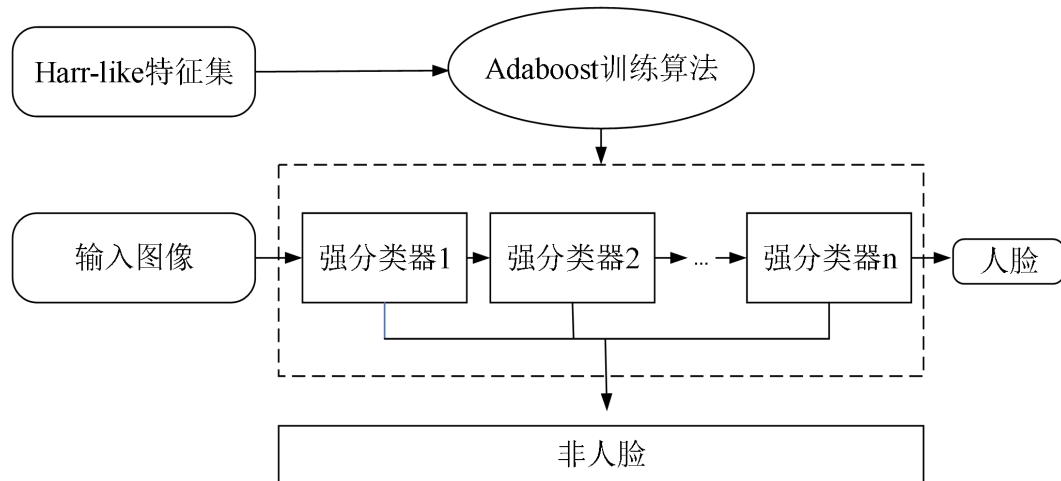


图 4-6 Viola-Jones 算法流程图

目标的面部包括很多特征，然而在一张正脸图像中各区域的特征的强弱以及特性都是不同的，例如，双眼部位比双颊部位暗，嘴唇区域也会比其他区域暗，这些特性存在于每个人的脸部，这些又可以通过黑白两种颜色的矩阵特征表示。Haar-like 特征由黑白的矩形特征组成，计算两种矩形的灰度差，通过它们来表示面部区域存在的特定特征。基础的是以下几种特征图，如图 4-7 所示，它们可以表示不同的特征模板，由它们来表征面部不同的特征。其中下面四种不同的积分图特征包括：边缘，线性、中心以及对角。

Viola-jones 人脸检测算法是一种基于滑动窗口的目标检测算法，但它却克服了滑动窗口检测带来的低效问题，可以用于实时人脸检测，利用一种新的数据结构积分图像来快速计算 Haar-like 特征。积分图像是受卷积的微分性质启发而定义一种数据结构。首先构造积分图，积分图像定义：

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x' + y') \quad (4-7)$$

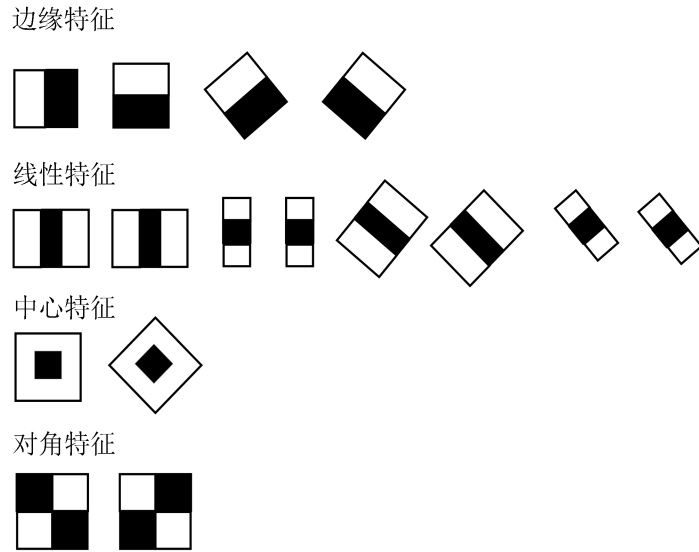


图 4-7 积分特征图分类

其中，积分图像任意点 (x, y) 位置处的像素值可以计算该点的左侧以及它的上方位置的所有像素值之和表示。通过积分图像之后将会快速计算得到 Haar-like 特征，如下图 4-8 所示。S1 到 S6 是积分图像在这六个顶点上的值。该特征值等于位于 A 中的像素总和减去位于 B 中的像素总和，而 A 中像素总和等于 $S5+S1-S2-S4$ ，B 中像素总和等于 $S6+S2-S3-S5$ ，并且无论矩形多大，我们总能在固定时间内计算出特征值（6 次索引操作和少量的加法乘法计算）。积分图像只需计算一次后续可以一直使用，事实上在算法实现时，我们只需保存样本的积分图像，原图像反而不用保存。

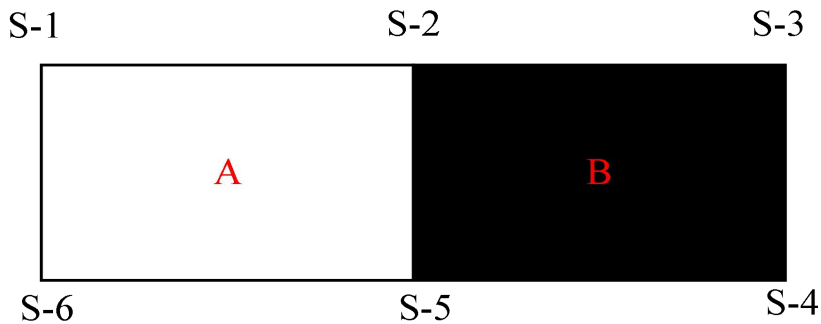


图 4-8 积分图特征图

特征选取步骤使用 AdaBoost 算法，每次训练的弱分类器的过程看作特征选取的过程，一次从 162336 个特征中选取一个特征（同时还包括了对应的门限值，极性，加权误差）。它的基本思想是训练一系列“弱”分类器，组成一个 committee（即每个弱分类器都有投票权，但是权重不同，加权误差越小的弱分类器权重越大）。adaboost 采用迭代训练方式，给定一个 t 阶 committee，如何寻找第 $t+1$ 个

弱分类器和对应的权重，以最小化在一定分布下的训练样本的加权指数损失。这个优化过程可以转换为对训练样本的分布来调整（即增大上一轮错误判断的样本的权重，减小正确判断的样本权重），在调整后的样本分布下寻找最小化加权 0-1 损失的弱分类器并计算对应的加权 0-1 损失。

4.3.2 色度特征提取

在进行心率检测时，脸部运动影响光线在皮肤表面进行的反射作用，这是在进行研究中不可避免的关键问题，很多研究者对整个脸部区域实施空间平均处理，这一操作旨在有效缓解表皮反射所导致的局部信号间断性变化问题，而血流体积的改变将会导致的持续变化累积起来。但是经过实验可知，因为目标移动等因素的影响，这个操作无法很好的解决这个问题。提取得到面部不同部位包含的信号，

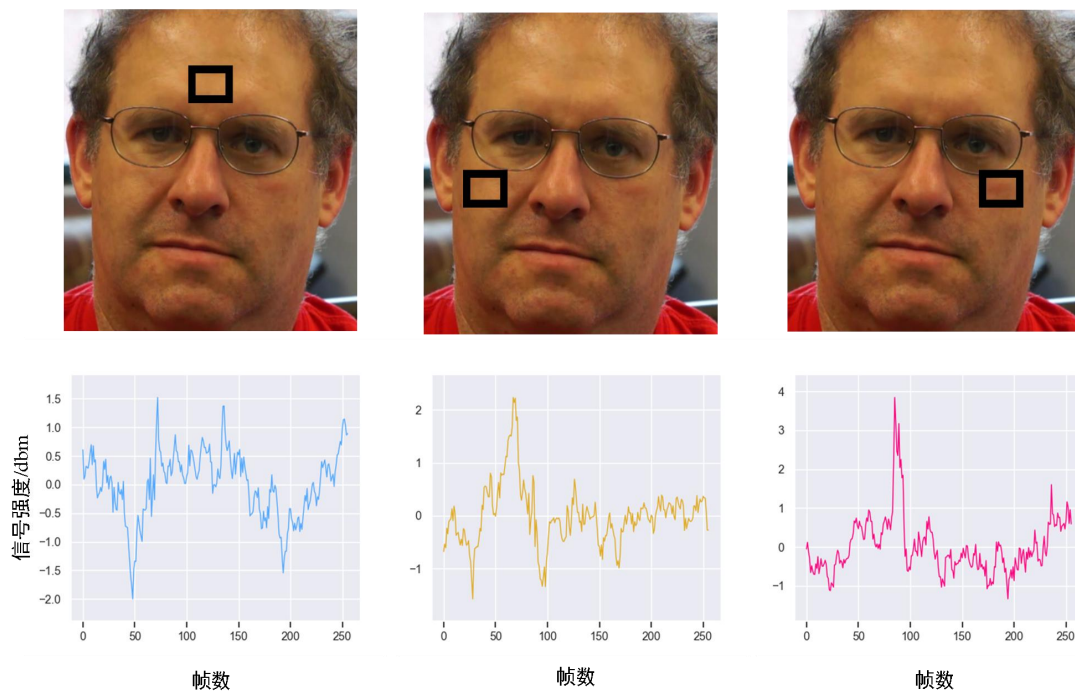


图 4-9 各区域的信号质量示意图

如图 4-9 所示。被测者额头，左脸颊，右脸颊信号质量明显不同。因此，简单的空间平均处理无法解决问题，还需进一步研究，选择合适的方法处理。

心率值的准确获取与得到的信号质量紧密相关。在本文方法中，选择色度特征进行脉搏波的提取。色度特征是基于皮肤反射模型的一种信号分解方式，它可以一定程度的去除因目标移动产生的噪声。得到的人脸视频中记录采集设备传感器中像素的反射光强度 $V(x, y, t)$ ，摄像机有三个颜色通道，能够分离这些通道，

利于后续信号处理，测得的 $V(x, y, t)$ ，分为两部分，入射光强度 $I(x, y, t)$ 和皮肤反射 $R(x, y, t)$ ，公式如下：

$$V(x, y, t) = I(x, y, t)R(x, y, t) \quad (4-8)$$

色度特征对于目标面部移动有较好的鲁棒性，它改变了估计投影方向，而色度特征可以减轻由于运动导致面部亮度以及颜色的变化。计算色度特征的公式如下：

$$C = X_f - \alpha Y_f \quad (4-9)$$

其中， $\alpha = \sigma(X_f) / \sigma(Y_f)$ ， σ 为标准偏差， X_f 、 Y_f 是 X, Y 使用带通滤波器后获得的信号。 $X = 3R_n - 2G_n$ ， $Y = 1.5R_n + G_n - 1.5B_n$ 。而 $R_n = R / \mu(R)$ ， $G_n = G / \mu(G)$ ， $B_n = B / \mu(B)$ 是三通道经过标准化后得到的信号。

4.3.3 信号加权与计算

通过预处理并且计算色度特征平均值之后，再进行一些简单的去噪操作，例如使用带通滤波操作，频率选择为 $[0.7, 4]\text{Hz}$ ，这个频率更适合得到纯净的心率信号，此时得到的是经过操作的粗心率信号。之前我们分析了面部区域各部分的信号质量都有所不同，在本文中我们要引入一个权值，将各区域信号与其权值进行融合，这样更有利于得到准确的心率值。图 4-10 为提出的基于 PSD 加权的脉搏波信号处理模块的流程图。具体步骤如下：

首先脉搏波信号中包括与心率信号相近的基础震荡频率，这个频率因其非常接近于心率信号。于是在确定心率信号的近似值时，我们选取功率谱中能量最为突出的频率作为参考：

$$PR = \arg \max_f |\hat{W}(f)| \quad (4-10)$$

$\hat{W}(f)$ 是 $\hat{u}(t)$ 在 $[0, T]$ 的 PSD，可以更精确地描述信号的功率分布情况。

在分析面部视频中的强度信号 $S(x, y, t)$ 时，首要步骤是对感兴趣区域内的每个像素区块提取其独特的色度特征 $u_i(t)$ 。接下来，针对所获得的色度特征 $u_i(t)$ ，采用频率范围限定在 $[0.5, 5.0]\text{Hz}$ 的带通滤波器进行处理，之后就会得到去除了部

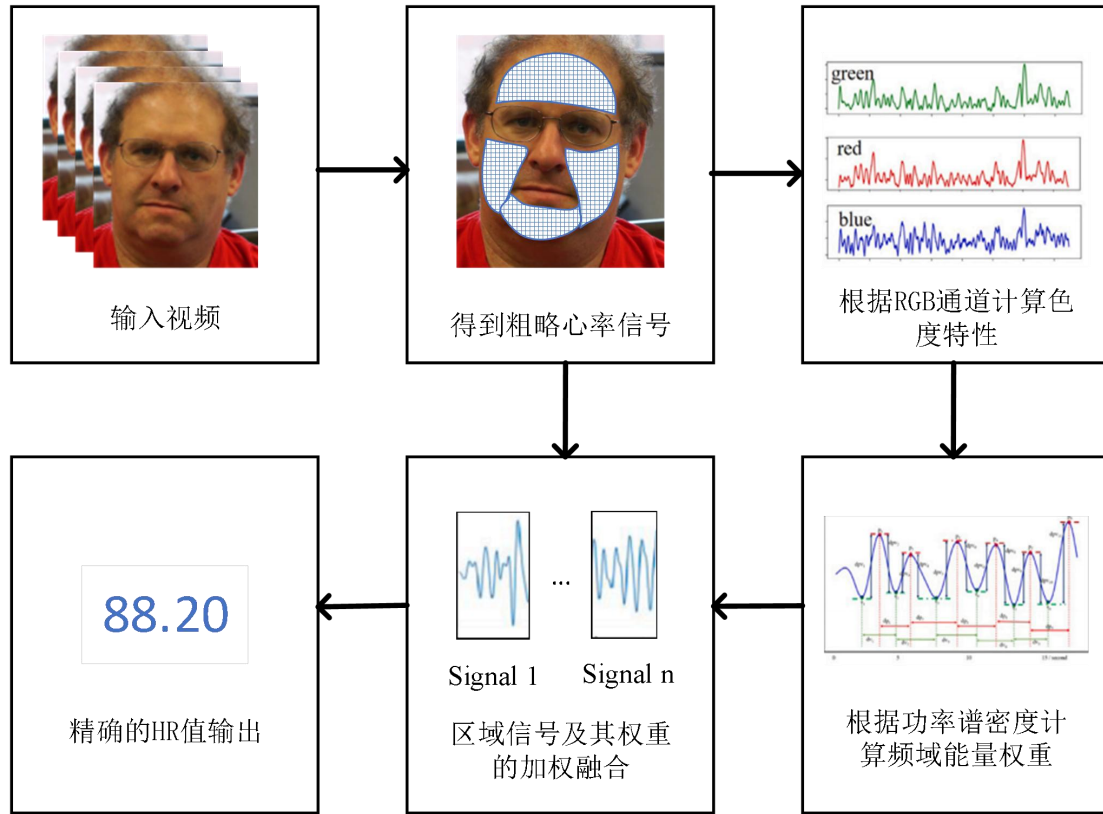


图 4-10 基于 PSD 加权的心率监测的流程

分噪声的信号，得到的信号为 $\hat{u}_i(t)$ 。图 4-11 左下图展示了未经过滤波处理前的心率信号，右下图是经过滤波器滤波后的心率信号。对比两幅图可以看出，使用滤波器之后的心率信号相比未使用前更加平滑，波形展示出明显的周期性。

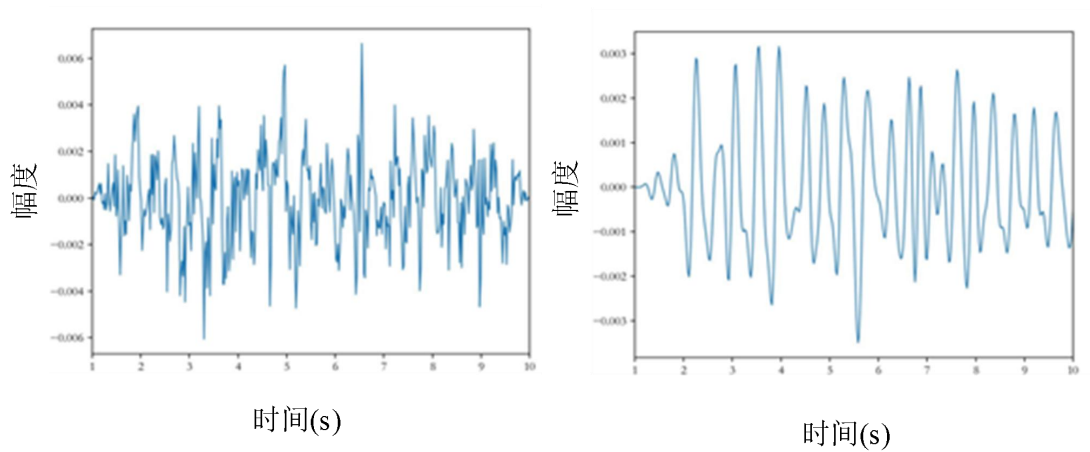


图 4-11 滤波处理心率信号示意图

将面部区域划分为 $k \times k$ 块，根据皮肤反射模型，各图像块使用滤波器后的信号为：

$$\begin{cases} \hat{u}_1(t) = B_1 p(t) + n_1(t) \\ \hat{u}_2(t) = B_2 p(t) + n_2(t) \\ \dots \\ \hat{u}_n(t) = B_n p(t) + n_n(t) \end{cases} \quad (4-11)$$

其中, $i \in \{1, 2, \dots, n\}$ 是感兴趣区域内的像素区块; B_i 代表 R_i 包含的 rPPG 强度, 该数值受到 α_i 和 I_i 的影响; $p(t)$ 代表血管内容积变化的信号; $n_i(t)$ 表示因运动伪影或者光照变化导致的噪声。之后将不同通道加权平均得到:

$$\hat{p}(t) = \sum_{i=1}^n w_i \hat{u}_i \quad (4-12)$$

各个通道所分配的权重 w_i 的设计借鉴了最大比合并算法的理念, 表明权重值与信号的均方根强度成正相关, 而与该通道自身的均方噪声强度成负相关, 这一规律可用如下数学公式表达:

$$w_i = \frac{A_i}{\|n_i(t)\|^2} \quad (4-13)$$

其中, 待估参数 A_i 和 n_i 均未知, 故需要借助其他估算策略来推算出 w_i 。值得注意的是, rPPG 信号的能量强度往往集中在一个与心率密切相关的频率区间, 而掺杂在心率信号中的背景噪声 $n_i(t)$ 则主要分布在 $[0.5, 5.0]$ Hz 的带通滤波器允许通过的频带范围外。据此, 可以通过在脉搏信号主导的特定频率段与该频带内噪声的谱能量比值, 来近似估算出 w_i 参数。

$$w_i = \frac{\int_{PR-h}^{PR+h} \hat{W}_i(f) df}{\int_{B_1}^{B_2} \hat{W}_i(f) df - \int_{PR-h}^{PR+h} \hat{W}_i(f) df} \quad (4-14)$$

其中, $\hat{W}(f)$ 为 $\hat{u}(t)$ 在 $[0, T]$ 时间段的 PSD, 某个特定频率段与心率脉搏信号紧密相关, 用 $[PR+h, PR-h]$ 表示; $[B_1, B_2]$ 是带通滤波器的工作频率范围。由上式可知, 信噪比高的区域所拥有的权重较大, 信噪比低的区域占有的权重较小。由此, 可以得到最终的精确的 rPPG 信号为:

$$V(t) = \sum_{i=1}^n w_i \hat{u}_i(t) \quad (4-15)$$

对精估计的 $rPPG$ 信号使用快速傅里叶变(FFT)分析功率谱如图 4-12 所示。最大功率对应的频率为 f_{HR} ，则估计的心率值 HR 为

$$HR = 60f_{HR} \quad (4-16)$$

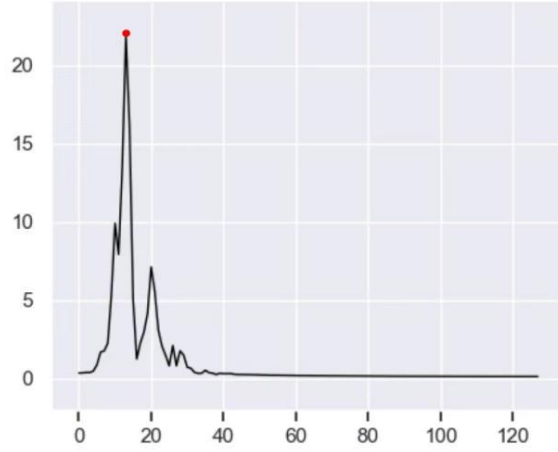


图 4-12 功率谱示意图

4.4 基于输入优化的 $rPPG$ 心率检测对比实验

本节将对对比各类输入优化算法的效果，验证选择的模型更适合本课题场景。我们选择传统的超分辨率算法以及基于深度学习的超分辨率算法进行比较，(1) 无上采样，(2) 双线性插值，(3) 双边插值(4) Real-Esrgan。传统的算法所有的分析都是在 MATLAB 中进行的，双线性和双边插值方法都是在计算机视觉工具箱中实现的，Real-Esrgan 算法使用 Python 实现。

使用 MAHNOB-HCI 数据集，从视频的不重叠的窗口中提取了 HR 估计值。这就产生了 N 个观察窗口。选择使用 MAE、RMSE 作为模型的客观性能指标。平均绝对误差 (MAE)：

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N} \quad (4-17)$$

均方根误差 (RMSE)：

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (4-18)$$

其中, y_i 表示真实值, $\hat{y}_i$ 表示预测值。我们展示了使用每种方法的结果。表 4-1 显示了每种方法的 MAE、RMSE。使用原始的高分辨率（高分辨）帧获得的结果也显示。

表 4-1 超分辨率算法模型性能对比

算法	MAE	RMSE
无上采样	4.41	6.79
双线性插值	3.56	5.45
双边插值	3.30	5.12
Real-Esrgan	2.86	3.98

由图表可知, Real-Esrgan 模型与其他传统超分辨率算法模型相比, 准确性较高, 因此本文选择 Real-Esrgan 模型作为预处理步骤使用模型。

4.5 基于 PSD 加权色度特征的心率估计验证实验

为了评估心率监测模块新算法的性能, 选取了 MAHNOB-HCI 数据集作为基准, 并与一系列已发表的传统方法文献[20], 文献[21], 文献[23], 文献[24], 文献[28]在精确度和运算速度两个维度上进行了详尽对比。在实验评估中使用数据集中的 EXG2 通道的信号作为真实心电信号。结果如图 4-13 所示, 文献[28]的准确率在比较算法中具有较强的优势, 但是速度仅有 9.7FPS, 速度较低。鉴于此, 我们借鉴了 Chrom 算法所采用的颜色空间投影分解技术, 这一策略有望在确保高准确度的同时, 显著提升算法运行速度, 从而实现更优的性能表现。

为了更加直观的看到本模块算法的性能, 绘制了图 4-14, 如图所示, 该图表示使用标准数据集测量得到的预测心率值以及对应的真实心率值, 可以看出本算法较为精准的预测得人体的心率值, 预测结果与真实结果较为接近, 具有很强的相关性。

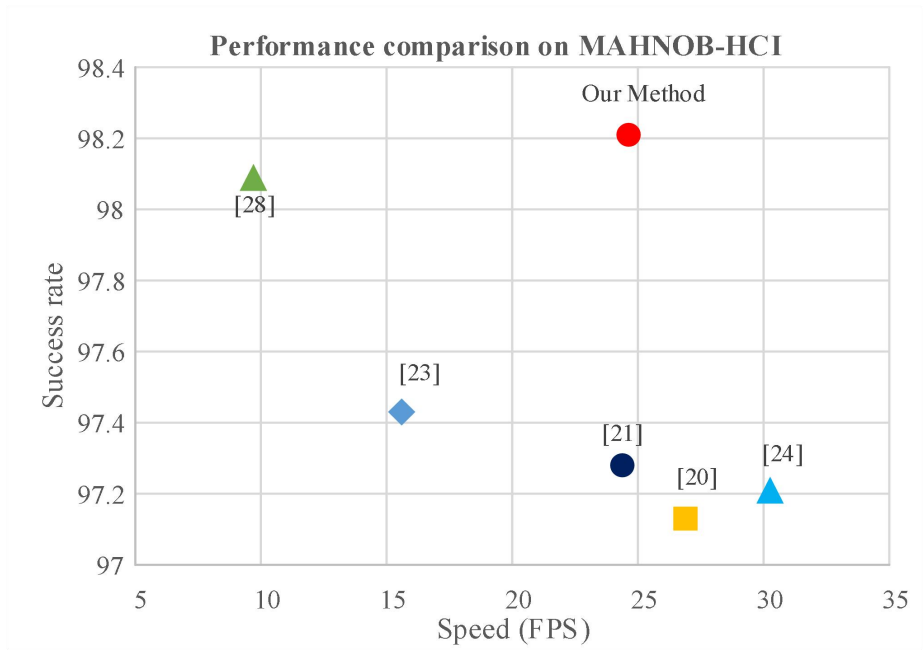


图 4-13 性能对比图

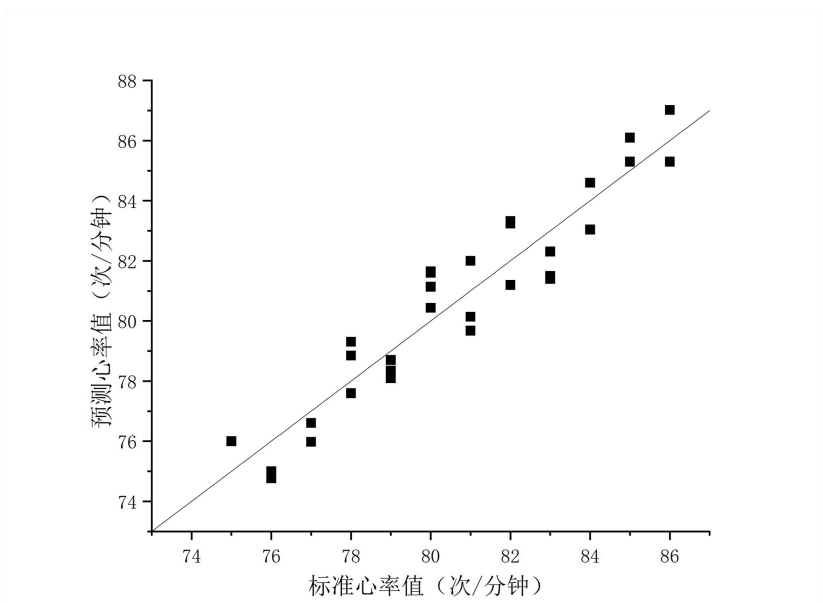


图 4-14 预测 HR 和标准接触 HR 的散点图

4.6 本章小结

本章提出了一种基于 PSD 加权色度特征的心率估计方法，首先进行视频预处理步骤，我们对比了几种传统的超分辨率方法的性能，选择合适的分辨率方法对于原始视频中丢失的面部信号进行恢复。之后选择色度特征并使用面部各区域信号加权的方法来消除因目标移动带来的干扰，在一定程度上改善了运动伪影带

来的影响，得到较为准确的心率值。最后进行多组实验，与传统算法比较，验证了方法的有效性。

第 5 章 实验与结果分析

5.1 本章引言

前三章探讨了移动目标心率监测算法的多个方面,结合考虑了噪声消除效果及算法的复杂性等多个关键因素,提出移动目标监测算法框架,包括第一模块的移动目标检测模块,以及引入超分辨率预处理步骤的基于 PSD 加权的心率监测模块。本节将移动目标心率监测算法应用于多个真实场景数据集中,与传统算法进行性能比较。

5.2 实验数据来源

实验所用的数据集包含了两部分内容:一部分为自主采集的数据集,另一部分则来源于公开在线资源 MAHNOB-HCI 数据集。通常情况下,自主收集的数据集被用来增强样本多样性,同时选用这些数据集与其他研究文献中已提出的传统算法成果进行对比验证。

MAHNOB-HCI 数据集是研究心率检测算法时常用的公开数据集。在数据集中包含 30 位不同肤色、不同国家、不同性别的健康成年被测者的采集信息。年龄范围涉及 19 岁到 40 岁(均值为 26.06,标准差为 4.39)。在数据的采集中,采集镜头包含一个处于受测者正前方的彩色镜头和五个处于不同方位的单色镜头。这些镜头将记录受测者的面部表情和头部姿势。实验将相关仪器连接到被测者各部位,并要求被测者跟随屏幕上的一个红圈来标定眼睛追踪器。实验装置以及数据集的示意图如图 5-1 所示。

在数据集中,包括心率,呼吸,血压等生理信息,如下表 5-1 所示。这些信号以 Biosemi data format(BDF)的格式存储。本文选择使用 EDFBrowser 读取信息。本文只使用数据集的 ECG 信息,所以只获取 EXG2 通道的数据。

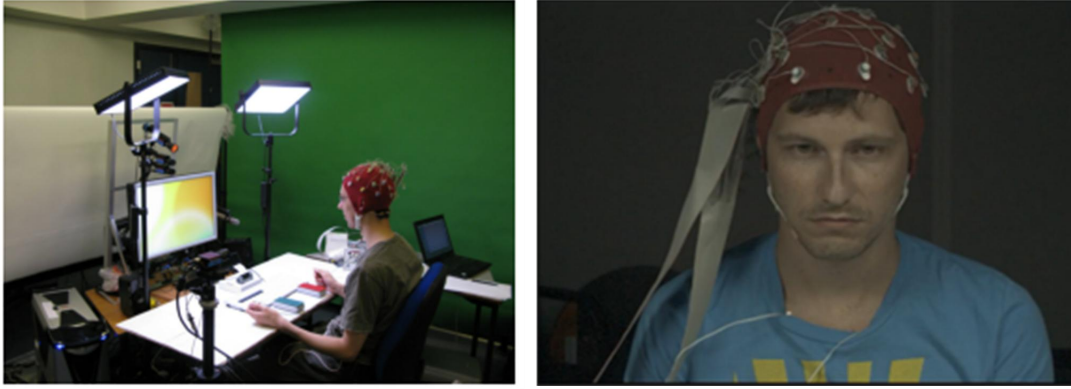


图 5-1 MAHNOB-HCI 数据集

表 5-1 MAHNOB-HCI 数据信息

Ch.no.	Ch.name	position	Physical unit
33	EXG1	ECG1(Upper right corner of chest,un-der clavicle bone)	uV
34	EXG2	ECG2(Upper left corner of chest,un-der clavicle bone)	uV
35	EXG3	Unused	uV
36	EXG4	Unused	-
37	EXG5	Unused	-
38	EXG6	Unused	-
39	EXG7	Unused	-
40	EXG8	Unused	-
41	GSR1	Galvanic skin response,left middle and ring finger	Ohm
42	GSR2	Unused	-
43	Erg1	Unused	-
44	Erg2	Unused	-
45	Resp	Respiration belt	uV
46	Temp	Temperature,left pinky	Celsius
47	Status	Status channel containing markers	Boolean

原创数据集采集方案概述如下：选取在慢速左右移动状态下的个体作为研究样本，只进行左右移动是为了保证距离摄像头的距离不变，排除该距离对实验结果带来不利影响，同时在左手食指上佩戴脉搏波采集设备（中科优瑞 UR-60 型），以测量数据作为真实 HR。手指式血氧仪使用光敏探测器采集数据，采集的数据包括血管内血容量变化情况，通过血管内反射出光的吸收量来得到心率值。光束经反射到达光电传感器，反射出的光中包含容积脉搏波，其中包含有效心率信息。血液容积变化情况反映出心脏跳动情况，最后将得到的数据信号经过电子电路和

微处理器处理后可以得到被测者的真实心率值，该仪器利用的是 PPG 测量心率的原理，最终心率值会显示在 LED 上。

自行收集的数据集是普通成年人的数据。被测目标一共有 5 个，年龄 20 到 30 周岁，女性 2 人，男性 3 人，每个目标有 8 个视频，时长为 30 秒，在实验室灯光照射下拍摄完成。视频拍摄过程中，每个测试者的视频片段如下：测试者直立面对摄像头 0.5、1、1.5、2m，没有动作或面部表情的变化。测试者面对摄像头 0.5、1、1.5、2m 进行左右移动，如图 5-2 所示。



图 5-2 自行采集数据集示意图

5.3 实验评估指标选择

本次实验采用了三种关键评估指标以对比和分析预测心率与实际心率间的差异，并评估预测心率的精确程度，这三项指标分别为：平均绝对误差、均方根误差以及 Bland-Altman 图。

5.3.1 误差指标

平均绝对值误差(MAE)，它表示预测值和真实值之间绝对误差的平均值。

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N} \quad (5-1)$$

均方根误差(RMSE)，它体现了预测值与真实值之间差异的平均样本偏差程

度。均方根误差这个指标代表了样本的离散程度。

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (5-2)$$

其中 N 表示样本数量, y_i 表示真实值, $\hat{y}_i$ 表示预测值。

5.3.2 Bland-Altman 图

Bland-Altman 图一般是用来评价两种连续变量测量方法的一致性, 一个是研究者提出的新方法, 另外一个公认的标准方法。对于同一个研究目标使用两种方法检测, 得到的结果在同一个维度, 可以进行比较。它能够直观的反映算法一致性和差异性。

5.4 实验结果分析

本次实验将使用自采集数据库进行测试, 效果如图 5-3 所示。与成像器的距离、被测者所处状态以及实验环境的光照强度作为本次实验的变量, 以稳定时间 30s 内的心率为有效值作为实验数据。使用文献[23]、文献[28]以及本文提出的算法分别对数据集进行心率监测, 共获得 2400 个心率数据。文献[23]提出的是一种基于独立分量分析(ICA)的心率监测算法, 文献[28]使用的是一种针对运动干扰的心率监测算法, 下文简称有界卡尔曼滤波方法, 它通过改进运动估计和特征跟踪的有界卡尔曼滤波技术, 来最小化由头部运动引起的模糊和噪声等运动伪影, 在一定程度上解决了目标移动带来的问题。用这两种传统算法与本文算法进行比较, 可以很好的验证本文算法在静止以及运动状态得到的结果的准确性。

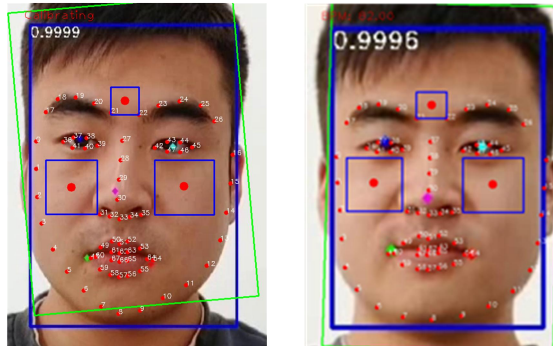


图 5-3 效果图

5.4.1 目标处于静止状态下的性能对比实验

本节将在目标处于静止状态下, 以距离成像器距离为变量, 光照强度为

100lux，采集视频分辨率为 1080p 的情况下，进行性能比较实验。表 5-2 可以看出，除了有界卡尔曼滤波方法以外另外两种方法的 HR(%)均高于 98%，并且本文以及 ICA 的 MAE 均在 1 左右，有较好的鲁棒性。由这些数据可知：当目标处于静止状态，且距离成像器 0.5m 时，是大多数研究者设定的研究场景，此时传统算法与本文算法的性能都较好，本文算法与其他两者传统算法性能差别不大。

表 5-2 目标静止状态，距离 0.5m 性能对比

算法	MAE	RMSE	HR(%)
[23]	1.08	1.19	98.64
[28]	1.32	1.59	98.31
本文	1.03	1.11	98.77

表 5-3 显示了目标处于静止状态下，距离成像器 1m 时的实验结果，从表 5-3 可以看出，此时的 MAE，RMSE 均有所提升，此时 HR(%)相比于 0.5m 距离时，整体大约下降不到 1%。由这些数据可知：该场景下，三种方法的准确度都有一定程度的下降，而本文算法优于其他两种传统算法。

表 5-3 目标静止状态，距离 1m 性能对比

算法	MAE	RMSE	HR(%)
[23]	1.22	1.61	98.18
[28]	1.45	1.87	97.91
本文	1.13	1.55	98.50

表 5-4 显示了目标处于静止状态下，距离成像器 1.5m 时的实验结果，从表可以看出，HR(%)数值只有本文算法大于 96%，而 RMSE 均高于 3，由此可知，三种算法性能均有所下降，其中，有界卡尔曼滤波方法的性能下降幅度最大。

表 5-4 目标静止状态，距离 1.5m 性能对比

算法	MAE	RMSE	HR(%)
[23]	2.89	3.75	95.73
[28]	3.16	3.99	95.41
本文	2.83	3.43	96.17

表 5-5 显示了目标处于静止状态下，距离成像器 2m 时的实验结果，三种算法性能均有所下降。此时与成像器的距离较远，大多数研究者都未曾研究过该距离的心率监测情况，因为远距离可以利用的生理信息较少，无法准确获得较为准确的心率估计值，然而 ICA 方法将三通道经过重新组合，有针对性的提升测量精度，而有界卡尔曼滤波方法重点则在抗运动场景下，此时该算法 HR(%)仅有 9

3.02%，性能最差。本文提出的算法在远距离静止场景下的性能较好，该算法添加了移动目标检测模块，可以有效的定位感兴趣区域，并通过预处理步骤，一定程度的降低了因为远距离带来的影响，所以本文在三者中性能最好。

表 5-5 目标静止状态，距离 2m 性能对比

算法	MAE	RMSE	HR(%)
[23]	3.27	4.80	94.53
[28]	5.64	6.18	93.02
本文	3.13	4.26	95.08

5.4.2 目标处于运动状态下的性能对比实验

本小节在目标处于运动状态下，以距离成像器距离为变量，光照强度为 100lux，采集视频分辨率为 1080p 的情况下，进行性能比较实验。表 5-6 显示了目标处于移动状态下，距离成像器 0.5m 实验结果。此时的距离是大多研究抗运动干扰算法的常见距离。三种算法的 RMSE 均高于 2，与静止状态相比，性能有所下降，这是由于目标移动带来的运动干扰导致的，但是准确率仍均大于 97%，说明，三种算法都有一定的抗干扰能力。其中，本文算法略优于另外两者算法。

表 5-6 目标运动状态，距离 0.5m 性能对比

算法	MAE	RMSE	HR(%)
[23]	1.58	2.79	97.11
[28]	2.82	2.41	97.62
本文	1.78	2.33	97.84

表 5-7 显示了目标处于移动状态下，距离成像器 1m 的实验结果。目标距离成像器距离有所增加，此时，三种算法中，只有本文提出的算法的准确率高于 96%，其中，ICA 算法受到运动干扰影响最大，RMSE 高于 4，抗运动干扰能力较弱。

表 5-7 目标运动状态，距离 1m 性能对比

算法	MAE	RMSE	HR(%)
[23]	3.53	4.09	95.16
[28]	2.68	3.57	95.99
本文	2.47	3.15	96.43

表 5-8 显示了目标处于移动状态下，距离成像器 1.5m 的实验结果。此时，三种算法的性能都有明显下降，其中，ICA 仍然是下降幅度最大的，此时它的准确率仅有 91.71%，此时有界卡尔曼滤波方法由于采用了基于运动估计和特征跟

踪的有界卡尔曼滤波技术，一定程度上最小化目标移动带来的干扰，因此，本文算法与有界卡尔曼滤波方法性能较为相似，本文算法略优于另一种算法。

表 5-8 目标运动状态，距离 1.5m 性能对比

算法	MAE	RMSE	HR(%)
[23]	6.91	7.19	91.71
[28]	5.24	5.68	93.60
本文	4.13	5.04	94.15

表 5-9 显示了目标处于移动状态下，距离成像器 2m 的实验结果。此时 ICA 的准确率已经低于 90%，有界卡尔曼滤波方法的准确率也仅有 91.23%。这个与成像器的距离是大多数研究者都未涉及的研究距离。此时，由于目标运动且距离成像器较远，感兴趣区域不仅较为模糊，而且受干扰影响，信号中包含的噪声信号也较多，需要的心率信号被大多数噪声信号掩盖，难以提取。本文添加的移动目标检测模块和超分辨率重建预处理，可以有效地定位感兴趣区域，一定程度的降低相机与目标之间的相对运动，并且恢复部分视频中丢失的生理信号，抗运动能力以及抗远距离能力都优于其他两者传统算法。

表 5-9 目标运动状态，距离 2m 性能对比

算法	MAE	RMSE	HR(%)
[23]	9.28	10.04	89.32
[28]	7.89	8.57	91.23
本文	5.63	6.12	93.17

以上实验可以证明在目标处于远距离移动状态下，本文的抗干扰能力优于之前提出的针对精度以及运动干扰而改进的两种传统算法。当目标处于静止状态的结果明显好于运动状态下的结果，这是由于运动带来较多噪声导致的。在距离成像器 2m 且被测者处于移动状态时，ICA 效果最差，而本文的算法优势将大大体现。进一步验证了本文提出的算法在目标处于移动状态且距离成像器较远距离场景下有较大优势，一定程度的提高了心率监测值的准确性。

5.4.3 目标处于不同光照强度下的性能对比试验

本文研究对象为移动目标，前两节研究了不同距离以及目标不同状态对算法的性能影响以及性能比较，本节将研究在目标与成像器距离相同，处于移动状态下，不同光照对算法的影响程度，并与真实平均心率值、文献[23]和文献[28]算法测得的平均心率估计值比较。为了排除被测者皮肤个体差异，我们将展示出数

据集中 5 个被测者的实验数据，结果如表 5-10 所示。

表 5-10 不光照强度下性能对比

实验对象编号	光强/ Lux	标准值	文献[23]	文献[28]	本文算法
1	30	67	74.88	73.15	71.84
	100	69	74.1	72.22	71.61
2	30	83	90.56	76.02	88.73
	100	83	88.18	79.01	86.41
3	30	88	95.12	82.01	93.11
	100	90	94.98	85.68	93.95
4	30	97	106.02	88.22	89.34
	100	98	102.13	94.14	94.28
5	30	69	76.09	75.13	74.42
	100	69	74.38	73.55	73.52

由表格表 5-10 可以看出三种算法在较弱光照强度下的性能都有所降低，30 Lux 光照强度较弱，在使用光电容及脉搏波方法进行基于面部视频的心率估计时，照射到皮肤的光照较弱，经过皮下血管介质后反射出的光照强度同样也较弱，不能很好的准确反映出血液容积变化情况，丢失较多的可利用的生理信息，再加上目标移动，光线与皮肤的角度发生变化，更容易产生噪声信号。由此可见合适光照条件也会对心率估计的结果带来不小的影响。

5.4.4 一致性分析实验

心率属于具有连续性的生理信号，需要客观衡量算法的心率估计值和真实心率值的相关性，为了更加直观地表现所提出算法监测心率的准确性，使用 Bland-Altman 图对本文所提出算法与使用设备得到的真实心率数据差异进行随机效应分析。实线表示均值，两条虚线之间的区域为 96%的一致性区域，可靠的数据点位于这个范围内。

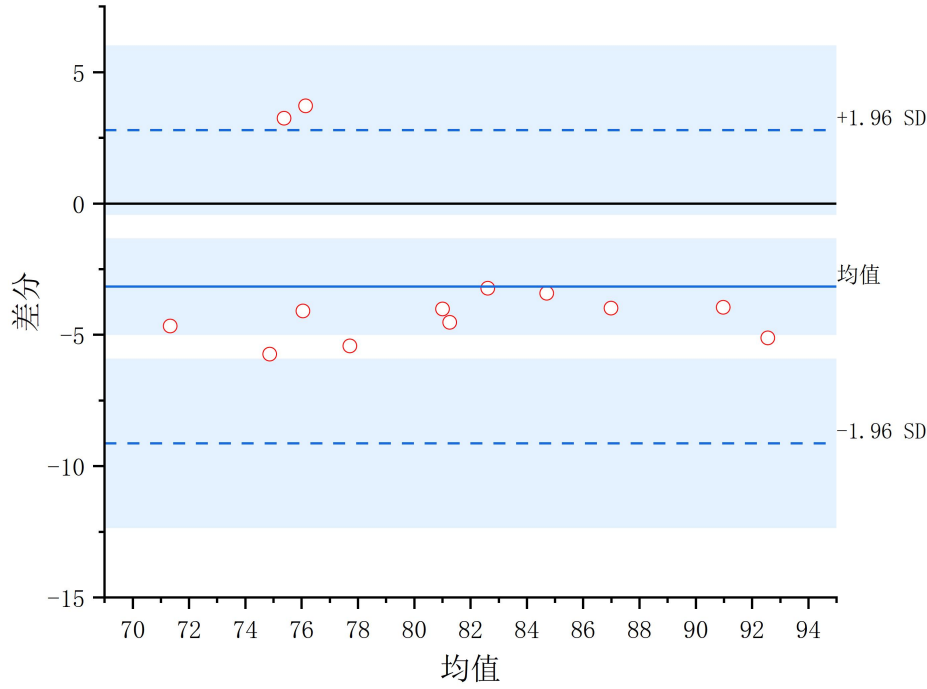


图 5-4 ICA 方法的 Bland-Altma 图

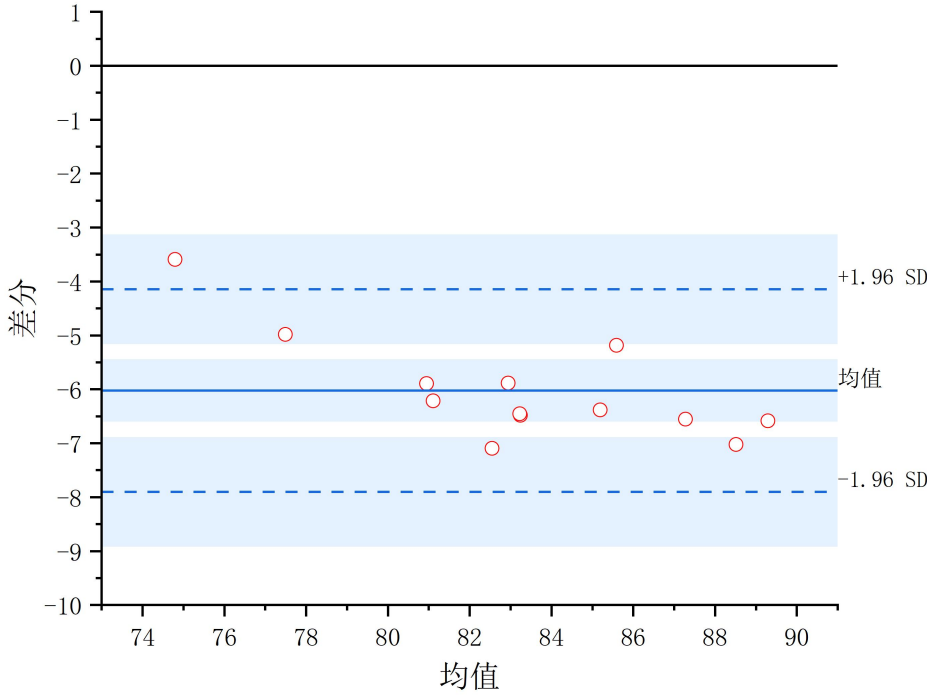


图 5-5 有界卡尔曼滤波方法的 Bland-Altman 图

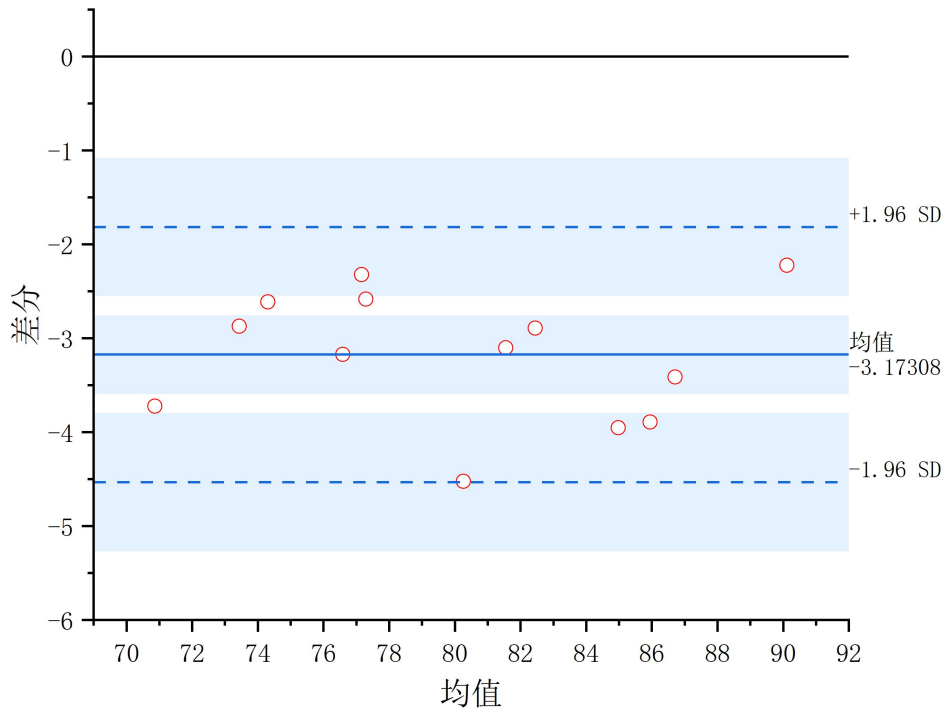


图 5-6 本文算法的 Bland-Altman 图

图 5-4，图 5-5，图 5-6 分别为 ICA 算法，有界卡尔曼滤波方法以及本文提出的算法在目标与成像器距离相同，光照条件相同，目标处于移动状态的情况下得到的 Bland-Altman 图。分析 Bland-Altman 图可知，图 5-4 有两个点位于置信区域外，这说明该算法的一致性较差，真实值与估计值匹配度较低，心率估计较为失败。图 5-5 只有一个点在置信区域外，有较好的一致性。图 5-6 是本文提出的算法，图中的点都位于置信区域内，说明本文算法得到的心率值没有出现较大偏差，测量结果具有良好的一致性。

5.5 本章小结

在被测者处于移动状态时，远程心率估计中的挑战较大。针对干扰因素，我们在第二章提出了移动目标心率监测框架，我们在传统框架内加入了移动目标检测模块，可以更好的定位被测者的位置，一定程度的降低移动带来的干扰；引入了超分辨率预处理步骤，提出基于 PSD 加权的脉搏波提取方法。在本章，我们使用公开和自行采集的数据集完成了一系列的实验，将本文算法与传统算法在不

同的场景下进行性能比较,证明了本文提出的算法在被测者处于移动状态的远程场景下具有不错的表现,证实了课题的可行性。

第 6 章 总结和展望

6.1 工作总结

本课题研究的移动目标的心率监测算法，包括两个模块。第一个模块是移动目标的检测模块，其目的是解决移动目标因移动不易被心率监测模型检测到并且检测到的效果较差的问题，方便后续心率监测的进行；第二模块是心率监测模块，对第一模块得到的视频进行预处理操作并引入 PSD 脉搏波信号处理方式，进而得到移动目标的心率值。本课题的研究成果在医学，体育等领域具有非常实用的价值，例如，在运动员运动时监测其心率值，给不能使用接触式心率监测仪器的病人监测其心率等等。

本文的主要研究工作如下：

（1）综合调研了国内外非接触式心率监测、移动目标检测等领域的研究现状。之后介绍了移动目标心率监测的应用前景以及出现的挑战，分析了目前常用的一些基于面部视频的心率检测方法以及基于深度学习的心率检测方法，分别介绍了它们各自存在的问题以及它们的优点，引出了本文的研究目的。其中借鉴 Huy Nguyen Quoc 的思想，该文介绍了一种基于 YOLOv3 检测器的新型人耳检测方法，在模型中添加了一个视网膜面部检测器，以缩小感兴趣的区域，提高精度。这个思想有利于解决我们在移动目标心率监测时出现的因目标移动对感兴趣区域不易被捕捉检测的问题。

（2）本文介绍了心率监测所使用的基础技术，分析得到影响该技术鲁棒性的一些干扰因素。复现并分析了现在大多数研究学者使用的心率监测框架，发现了基于视觉的心率监测框架应用于目标移动的场景出现的一些问题。针对这些问题，提出了一种新的基于移动目标的心率监测框架。介绍了基于移动目标的心率监测框架的各模块的作用以及意义。

（3）针对基于移动目标的心率监测框架的第一模块，介绍了使用的基础模型 YOLOv4，并针对该模型应用于本课题出现的检测模型参数量较大，实时性较差，远距离检测移动目标导致小尺度目标关注度变低，存在遮挡情况进而出现目

标特征语义信息不足,检测精度降低等问题,提出基于 YOLOv4 轻量化的移动目标检测模型。引入 GhostNet 替换原始的 CSPDarknet-53 模块,在模型非主干部分后面添加 ECANet 机制,提高模型的检测精度降低模型参数量。使用经过筛选的 VOC2007 数据集,与其他轻量级模型进行对比,证明经过改进的模型在一定程度上提升了速度以及精度,为之后的心率监测模块奠定一定的基础。

(4) 针对移动目标心率监测框架的第二模块,介绍了一些输入优化的模型,经过对比实验选择使用 Real-ESRGAN 模型进行在心率监测之前的预处理步骤,介绍了该模型的基础结构,并介绍该预处理步骤的目的,即解决因距离成像器较远或者输入视频有所压缩导致的生理信号的丢失的问题,一定程度的恢复了视频丢失的信息,有利于提高后续心率监测的准确性。针对目标移动产生的运动伪影,选择色度特征信号使用面部各区域信号加权的方法来消除现实环境中脸部运动的干扰,提取准确的脉搏波信号波形。介绍了皮肤反射模型和色度特征的提取以及提出的基于 PSD 加权的心率估计方法。使用 MAHNOB-HCI 数据集进行验证,将该方法与 ICA、Green、Chrom 等其他传统方法进行比较。经过验证发现该方法在精度与速度方面都有不错的表现。

(5) 针对本课题提出的移动目标心率监测方法,使用公开和自行采集的数据集完成一系列的验证实验。介绍实验用到的评价指标以及实验用到的数据集。在不同的场景下,将本文方法与传统方法进行性能比较,分析实验结果,得出结论。经过一系列实验证明了本文提出的算法在被测者处于移动状态的远程场景下具有不错的表现,证实了课题的可行性。

6.2 展望

本论文对于移动目标心率监测方法进行研究,提出一种基于移动目标心率监测框架,对于各模块进行改进。但是仍有部分工作需要在未来开展,部分想法需要进一步验证并实现,存在的一些不足以及需要进一步研究的方向如下:

(1) 经过实验可知随着目标距离成像器的距离越远,得到的心率值的准确度越低,虽然本文已在较远距离的准确性方面有一定的突破,但是在之后的研究中需要进一步改进以应对更远的距离以及更剧烈的移动状态。

(2) 本课题自行采集进行验证的数据集人数相对较少,收集的被测者皆为

身体健康的青年人。在之后的研究中，将会收集年龄范围更广，不同肤色及国籍，并非全是身体健康的被测者的心率数据，增加更多实验变量，研究不同因素对于心率值的影响情况。提出更加全面实用的方法应用于医学，体育领域。

参考文献

- [1] Lee K J, Kim B J, Han M K, et al. Effect of Heart Rate on 1-Year Outcome for Patients With Acute Ischemic Stroke[J]. Journal of the American Heart Association, 2022, 11(10): e025861.
- [2] Olshansky B, Ricci F, Fedorowski A. Importance of resting heart rate.[J].Trends Cardiovasc Med. 2022 ,33(8):502-515.
- [3] Baka T, Simko F. Monitoring non-dipping heart rate by consumer-grade wrist-worn devices: an avenue for cardiovascular risk assessment in hypertension[J]. Frontiers in Cardiovascular Medicine, 2021, 8: 711417.
- [4] Tawfello G, Gulko D, Petrossian G, et al. Importance of Resting Heart Rate Assessment in Autonomic Dysfunction Evaluation in Congestive Heart Failure Patients[J]. Circulation, 2023, 148(Suppl_1): A11820-A11820.
- [5] 刘士贺,韩旭,刘倩,等.静息心率对超高危动脉粥样硬化性心血管疾病患者全因死亡风险的影响[J].中国循环杂志,2024,39(02):140-147.
- [6] 张希铃,王新康.心率变异性在心血管疾病中的研究进展[J].实用心电学杂志,2023,32(05):365-369.
- [7] 李嘉华,蔡婧婧,任力杰.可穿戴医疗设备在卒中高危人群和患者中的应用[J].中国卒中杂志,2021,16(12):1290-1295.
- [8] An R. Comparative Study of Spo2 in all the Fingers of the Hands Measured by Pulse Oximeter[J]. The Journal of the Association of Physicians of India, 2022, 70(4): 11-12.
- [9] 舒丽瑾,许艳荣.手环心率监测下有氧运动在慢性心衰中的效果[J].新疆医学,2022,52(06):658-661.
- [10] 谢佳玲,龚渝顺,魏良,等.一种针对可穿戴设备的抗运动干扰心率检测算法[J].生物医学工程学杂志,2021,38(04):764-773.
- [11] Zanelli S, El Yacoubi M A, Hallab M, et al. Transfer learning of CNN-based signal quality assessment from clinical to non-clinical PPG signals[C]//2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2021: 902-905.

- [12]Jang, Seungju. Design of a Heart Rate Measurement System Using a Web Camera[J].The International Journal of Advanced Smart Convergence, 2022, 11(3): 179-186.
- [13]Ogawa K, Koyama S, Haseda Y, et al. Wireless, portable fiber Bragg grating interrogation system employing optical edge filter[J]. Sensors, 2019, 19(14): 3222.
- [14]Azzouz A, Bengherbia B, Wira P, et al. An efficient ECG signals denoising technique based on the combination of particle swarm optimisation and wavelet transform[J]. Heliyon, 2024, 10(5):e26171.
- [15]Yin C, Zhang Q, Ren W. A new semantic edge aware network for object affordance detection[J]. Journal of Intelligent & Robotic Systems, 2022, 104(1): 2.
- [16]宋正伟. 基于腕部桡动脉的脉搏波提取方法研究[D].合肥工业大学,2022.
- [17]Sohn J, Shin H, Lee J, et al. Validation of Electrocardiogram Based Photoplethysmogram Generated Using U-Net Based Generative Adversarial Networks[J]. Journal of Healthcare Informatics Research, 2024, 8(1): 140-157.
- [18]周娟, 李丹, 龙云铸. 新型冠状病毒(2019-nCoV)相关研究进展[J]. 中国感染控制杂志, 2020,03:288-292.
- [19]Guang-da L, Xiao-hu S, Xing W, et al. Determination of Circulation Blood Volume by Modified Pulse Dye Densitometry[J]. SPECTROSCOPY AND SPECTRAL ANALYSIS, 2018, 38(3): 727-731.
- [20]Verkruysse W, Svaasand L O, Nelson J S. Remote plethysmographic imaging using ambient light [J]. Optics Express, 2008, 16(26):21434-21445
- [21]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.
- [22]戴阳,郑婷婷,杨雪.基于视频放大与盲源分离的非接触式心率检测[J].计算机系统应用,2021,30(01):228-234.
- [23]Bhowmik M K, Saha P, Singha A, et al. Enhancement of robustness of face recognition system through reduced gaussianity in Log-ICA[J]. Expert Systems

- with Applications, 2019, 116: 96-107.
- [24]De Haan G, Jeanne V. Robust pulse rate from chrominance-based IPPG [J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2878-2886.
- [25]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.
- [26]Wang W, den Brinker A C, Stuijk S, et al. Amplitude-selective filtering for remote-PPG [J]. Biomedical optics express, 2017, 8(3): 1965-1980
- [27]Chwyl B, Chung A G, Amelard R, et al. SAPPHERE: Stochastically acquired photoplethysmogram for heart rate inference in realistic environments[C]//2016 IEEE international conference on image processing (ICIP). IEEE, 2016: 1230-1234.
- [28]Prakash S K A, Tucker C S. Bounded Kalman filter method for motion-robust, non-contact heart rate estimation [J]. Biomedical optics express, 2018, 9(2): 873-897.
- [29]Špetlík R, Franc V, Matas J. Visual heart rate estimation with convolutional neural network[C]//Proceedings of the british machine vision conference, Newcastle, UK. 2018: 3-6.
- [30]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.
- [31]Wang Z K, Kao Y, Hsu C T. Vision-based heart rate estimation via a two-stream cnn[C]//2019 IEEE international conference on image processing (ICIP). IEEE, 2019: 3327-3331.
- [32]Blackford E B, Estepp J R, Piasecki A M, et al. Long-range non-contact imaging photoplethysmography: cardiac pulse wave sensing at a distance[C]//Optical Diagnostics and Sensing XVI: Toward Point-of-Care Diagnostics. SPIE, 2016, 9715: 176-192.
- [33]Gao H, Zhang C, Pei S, et al. Region of Interest Analysis Using Delaunay Triangulation for Facial Video-Based Heart Rate Estimation[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-12.

- [34] Yu Z, Peng W, Li X, et al. Remote heart rate measurement from highly compressed facial videos: an end-to-end deep learning solution with video enhancement[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 151-160.
- [35] Unke O T, Meuwly M. PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges[J]. Journal of chemical theory and computation, 2019, 15(6): 3678-3693.
- [36] 王杨,赵红东.基于改进粒子群优化的支持向量机与情景感知的人体活动识别[J].计算机应用, 2020,40(3):51-57.
- [37] Pacal I, Karaman A, Karaboga D, et al. An efficient real-time colonic polyp detection with YOLO algorithms trained by using negative samples and large datasets[J]. Computers in biology and medicine, 2022, 141: 105031.
- [38] Xiao Y, Zhou K, Cui G, et al. Deep learning for occluded and multi-scale pedestrian detection: A review[J]. IET Image Processing, 2021, 15(2): 286-301.
- [39] 谢斌红,袁帅,龚大立.基于 RDB-YOLOv4 的煤矿井下有遮挡行人检测[J].计算机工程与应用, 2022, 58(05):200-207.
- [40] 康帅, 章坚武, 朱尊杰, 童国锋.改进 YOLOv4 算法的复杂视觉场景行人检测方法[J].电信科学, 2021, 37(08):46-56.
- [41] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). Ieee, 2005, 1: 886-893.
- [42] Yuan B H, Liu G H. Image retrieval based on gradient-structures histogram[J]. Neural Computing and Applications, 2020, 32: 11717-11727.
- [43] Li Y, Fan L, Xie W. TGSIFT: Robust SIFT descriptor based on tensor gradient for hyperspectral images[J]. Chinese Journal of Electronics, 2020, 29(5): 916-925.
- [44] Kim D, Lee Y, Park C G. Context awareness and step length estimation by shape distance and H-features[J]. International Journal of Control, Automation and Systems, 2020, 18(12): 3051-3061.
- [45] Yang J, Gao H. Cultural emperor penguin optimizer and its application for face recognition[J]. Mathematical Problems in Engineering, 2020: 1-16.

- [46]Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model[C]//2008 IEEE conference on computer vision and pattern recognition. Ieee, 2008: 1-8.
- [47]Ke X, Lin X, Qin L. Lightweight convolutional neural network-based pedestrian detection and re-identification in multiple scenarios[J]. Machine Vision and Applications, 2021, 32: 1-23.
- [48]Hussain M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection[J]. Machines, 2023, 11(7): 677.
- [49]Liu J, Zhu X, Zhou X, et al. Defect detection for metal base of TO-Can packaged laser diode based on improved YOLO algorithm[J]. Electronics, 2022, 11(10): 1561.
- [50]Wang M, Yang B, Wang X, et al. YOLO-T: multitarget intelligent recognition method for x-ray images based on the YOLO and transformer models[J]. Applied Sciences, 2022, 12(22): 11848.
- [51]Hurtik P, Molek V, Hula J, et al. Poly-YOLO: higher speed, more precise detection and instance segmentation for YOLOv3[J]. Neural Computing and Applications, 2022, 34(10): 8275-8290.
- [52]Zhai S, Shang D, Wang S, et al. DF-SSD: An improved SSD object detection algorithm based on DenseNet and feature fusion[J]. IEEE access, 2020, 8: 24344-24357.
- [53]Xu X, Zhao M, Shi P, et al. Crack detection and comparison study based on faster R-CNN and mask R-CNN[J]. Sensors, 2022, 22(3): 1215.
- [54]Purkait P , Zhao C , Zach C .SPP-Net: Deep Absolute Pose Regression with Synthetic Views[C]//Computer Vision and Pattern Recognition (CVPR) 2017:1-10.
- [55]Kızılluk S, Sert E. Hurricane-Faster R-CNN-JS: Hurricane detection with faster R-CNN using artificial Jellyfish Search (JS) optimizer[J]. Multimedia Tools and Applications, 2022, 81(26): 37981-37999.

- [56]Ren S , He K , Girshick R , et al. Faster R-CNN: Towards Real-Time , Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6):1137-1149.
- [57]Malekghasemi M, Bagheri S M. Consistency and interpolation in linear continuous logic[J]. Archive for Mathematical Logic, 2023, 62(7-8): 931-939.
- [58]NAYAK R, PATRA D. Image interpolation using adaptive P-spline[C]//Proceedings of the 2015 Annual IEEE India Conference, New Delhi, Dec 17- 20, 2015. Piscataway: IEEE, 2015: 1-6.
- [59]Aràndiga F. A nonlinear algorithm for monotone piecewise bicubic interpolation[J]. Applied Mathematics and Computation, 2016, 272: 100-113.
- [60]Agathiyan A, Gowrisankar A, Fataf N A A, et al. Remarks on the integral transform of non-linear fractal interpolation functions[J]. Chaos, Solitons & Fractals, 2023, 173: 113749.
- [61]刘克俭, 陈淼焱, 冯琦. 一种优化的迭代反投影超分辨率重建方法[J]. 遥感信息, 2019, 34(03):14-18.
- [62]Akrou R, Atamna A, Peters J. Convex optimization with an interpolation-based projection and its application to deep learning[J]. Machine Learning, 2021, 110: 2267-2289.
- [63]聂涛. 基于最大后验概率的视频超分辨率算法研究[D].浙江大学,2015.
- [64]Luo J, Liu L, Xu W, et al. Stereo super-resolution images detection based on multi-scale feature extraction and hierarchical feature fusion[J]. Gene Expression Patterns, 2022, 45: 119266.
- [65]Wang X, Hu Q, Cheng Y, et al. Hyperspectral image super-resolution meets deep learning: A survey and perspective[J]. IEEE/CAA Journal of Automatica Sinica, 2023, 10(8): 1668-1691.
- [66]Zhou C, Xiong A. Fast image super-resolution using particle swarm optimization-based convolutional neural networks[J]. Sensors, 2023, 23(4): 1923.
- [67]Freeman W T, Pasztor E C, Carmichael O T. Learning low-level vision[J]. International Journal of Computer Vision, 2000, 40: 25-47.

- [68]Chang H, Yeung D Y, Xiong Y. Super-resolution through neighbor embedding[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 2004: 275-282.
- [69]Gao X, Zhang K, Tao D, et al. Image super-resolution with sparse neighbor embedding[J]. IEEE Transactions on Image Processing, 2012, 21(7): 3194-3205.

攻读硕士期间发表学术成果

- (1) 以第一作者在《牡丹江师范学院学报》发表论文一篇。（已录用）
- (2) 以第一作者在 SCI 期刊《Optical Engineering》投稿论文一篇。（小修中）
- (3) 以导师第一，本人第二申请国家发明专利一项。（实审中）

致谢

凡是过往，皆为序章。随着这篇致谢之词的开始，意味着我在安徽工程大学三年的硕士学习时光在这 2024 年盛夏即将结束。时间是让人猝不及防的东西，晴时有风阴时有雨，目光所及之处，皆为回忆。纵然有千万般不舍，但仍然心存感激。

诚挚地感谢我的导师王力超副教授。王力超老师对学生要求严格，使我在三年的硕士学习中进步良多。自从入校以来，老师在科研、学习上对我悉心指导，在生活上也给予我无微不至的关怀和照顾。老师优秀的品质和精益求精的工作态度，对我有着深远的影响，使我受益终身。感谢王力超老师这三年不辞辛苦的指导我，从论文选题、实验设计、到期刊论文和毕业论文的撰写，都倾注了您的心血。最后在撰写毕业论文时，您不厌其烦的为论文提出修改意见，使我的论文撰写工作有条不紊地完成。因此在此论文即将完成之际，向王力超老师表示诚挚的感谢。

感谢研究生三年朝夕相处的同学们。这三年时间，我们一起经历了疫情，一起在实验室朝九晚十的认真科研，感谢同学们的相识，陪我走过了三年难忘的历程。感谢徐莎曼、尚永健、董家辉这三个充满正能量的同学们，因为你们让这三年生活不枯燥，每天都非常开心快乐。感谢已经毕业的学长学姐王义琼和李子阳，在学习和生活中给予我许多帮助。感谢我的学弟学妹们刘飞瑶、缪陆、刘俊楠、尹冰、汪冉、刘祥玉、黄志成。还要感谢我的舍友项诗雨、李田利、潘孟星同学，她们给宿舍带来了非常多的欢乐。祝同学们前程似锦，平安喜乐，顺顺利利。

感谢一直陪伴在我身边的父母。感谢我的父母不辞辛苦的养育我，支持我的每个决定，在生活上无微不至的照顾我，让我能够安心学习，给予我无私的爱。父母之恩，无以为报，我会用一生去爱你们，希望未来成为你们可以依靠的肩膀，在此感谢你们。

此刻，落笔至此。我的研究生三年学习生活交上了一份满意的答卷，同时也要感谢自己，感谢自己在遇到困难时没有轻易放弃。山水一程，三生有幸，以梦为马，不负韶华！

尚德敏学 唯实唯新

