

多部位皮肤检测研究及其在洗浴机器人中的应用

李平^{1,2}, 喻洪流¹

(1. 上海理工大学 智能康复工程研究院, 上海 200093; 2. 长治医学院 生物医学工程系, 长治 046000)

摘要: 为了应对人口老龄化带来的挑战, 并使洗浴机器人能够针对不同皮肤部位采用适当的洗浴模式, 对多部位皮肤检测及其在洗浴机器人中的应用进行了研究。在前期研究的基础上, 选取了4种典型目标检测算法, 扩充了原始数据集, 并基于迁移学习进行多部位皮肤检测。建立了综合评价指标以评估算法性能, 在Tesla T4和TX2平台上对性能最佳的模型进行部署和测试, 并将其应用于洗浴机器人中。结果显示: 数据集类不平衡的改善可使检测精度平均提升18%; YOLOv5s算法在精度与模型大小之间达到了最佳平衡, 能够在Tesla T4和TX2平台上进行实时检测, 并在水汽环境中实现对不同部位皮肤的识别。通过TX2平台集成视觉传感器, 进行目标点三维位姿建模和联合实验, 控制机器人到达背部区域的成功率为92%, 使用点云作为监督信息可将此成功率提升至100%。改善类不平衡可以显著提升多部位皮肤检测的准确性, YOLOv5s在平衡精度和模型大小方面表现出色, 有效满足了洗浴机器人多部位皮肤检测的需求。

关键词: 皮肤检测; 洗浴机器人; YOLOv5s算法; 迁移学习; TX2平台

中图分类号: TP 391.4 文献标志码: A

Multi-part skin detection and its application in bathing robots

LI Ping^{1,2}, YU Hongliu¹

(1. Institute of Intelligent Rehabilitation Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;

2. Department of Biomedical Engineering, Changzhi Medical College, Changzhi 046000, China)

Abstract: To address the challenges posed by an aging population and to enable bathing robots to apply appropriate cleaning modes for different skin regions, the research was conducted on multi-part skin detection and its application in these robots. Building on previous studies, four representative object detection algorithms were selected, the original dataset was expanded, and multi-part skin detection was

收稿日期: 2024-06-13

基金项目: 国家重点研发计划资助项目(2022YFC3601403); 国家自然科学基金资助项目(62073224); 山西省高等学校科技创新项目(2022L376); 长治医学院博士科研启动基金项目(2024BS12)

第一作者: 李平(1989-), 女, 讲师。研究方向: 计算机视觉、康复机器人和护理机器人。E-mail: lmxlp@126.com

通信作者: 喻洪流(1966-), 男, 教授。研究方向: 智能康复机器人、智能人机交互。E-mail: yhl98@hotmail.com

引文格式: 李平, 喻洪流. 多部位皮肤检测研究及其在洗浴机器人中的应用[J]. 上海理工大学学报, 2025, 47(1): 1-8.

Citation: LI Ping, YU Hongliu. Multi-part skin detection and its application in bathing robots[J]. Journal of University of Shanghai for Science and Technology, 2025, 47(1): 1-8.

performed using transfer learning. A comprehensive evaluation metric was developed to assess algorithm performance, and the best-performing model was deployed and tested on Tesla T4 and TX2 platforms, with practical applications in bathing robots. The results showed that addressing dataset class imbalance lead to an average accuracy improvement of 18%. YOLOv5s has achieved an optimal balance between accuracy and model size, facilitating real-time detection on both Tesla T4 and TX2, and effectively identifying different skin regions in a humid environment. By integrating a visual sensor with TX2, 3D pose modeling and joint experiments were conducted, achieving a 92% success rate in controlling the robot to reach the back area. Utilizing point cloud data as supervision improve the success rate to 100%. Addressing class imbalance significantly enhances the accuracy of multi-part skin detection, and YOLOv5s exhibits excellent performance in balancing accuracy and model size, effectively meeting the multi-part skin detection requirements of bathing robots.

Keywords: *skin detection; bathing robots; YOLOv5s algorithm; transfer learning; TX2 platform*

人口老龄化趋势给家庭和保险系统带来了巨大的经济负担,增加了对特殊护理的需求,特别是在如厕和洗澡等日常活动方面。洗浴辅助设备很多,但大部分设备功能单一,无法提供个性化、智能化服务,洗浴任务的完成仍主要依靠护理人员的辅助。护理人员不足、护理工作量大等问题,造成了用户和护理人员之间的尴尬及不可预知的一些安全风险(如传染病传染的潜在风险),洗浴机器人的研发成为解决上述问题的重要途径,并成为当前的研究热点^[1]。为了对不同部位的皮肤提供特定的清洗模式,洗浴机器人中的皮肤检测功能不仅需要区分皮肤与非皮肤,还必须识别出皮肤的具体部位(如背部、上肢等)。

在洗浴场景中进行皮肤检测时:首先,将检测所得信息转换至机器人坐标系;然后,通过运动规划得到末端执行器位姿,从而控制机器人向目标区域运行;最后,引导机器人到达目标皮肤区域完成洗浴任务(擦洗、冲洗等)。可见,皮肤检测是实现洗浴任务的关键环节之一。

根据检测原理,皮肤检测可分为3种方法:基于像素、基于区域和混合方法。基于像素的方法将图像中的每个像素分为皮肤和非皮肤^[2];基于区域的方法识别图像中具有相似特征的区域;混合方法使用不止一种检测原理^[3]。根据所使用的特征,皮肤检测可以大致分为基于颜色的^[4]、基于纹理的^[5]、基于统计特征^[6]和基于其他特征^[7]的方法。从是否明确建立皮肤模型的角度来看,皮肤检测可以分为基于机器学习的方法^[8]和传统的方法^[9],前者通常使用监督方法构建皮肤检测器,后

者一般明确地建立皮肤模型。

皮肤检测可用于黑色素瘤、皮肤癌等疾病的诊断^[10-11]。而在面向洗浴机器人的皮肤检测研究中,Zlatintsi等^[12]使用语义分割,为图像中的每个像素分配类别,但该分类结果对像素分类错误较为敏感,且不利于机器人的实时控制。使用目标检测进行多部位皮肤检测时,无需对每个像素进行判断来区分皮肤不同区域,以及为洗浴机器人提供控制信息。与语义分割相比,目标检测的标注工作量和计算工作量较小,有利于机器人的实时控制。在洗浴机器人中,针对人体不同部位皮肤设置不同的洗浴模式。皮肤检测不仅需要识别皮肤和非皮肤区域,还要检测皮肤属于人体的哪个部位以便进行洗浴模式选择,保障洗浴效果。机器人移动到不同部位皮肤后,可以执行预先设计的运动方式,如垂直擦洗、水平擦洗、划圈式擦洗等。

近年来,基于深度卷积神经网络(deep convolutional neural networks, DCNNs)的目标检测算法发展迅速,它以大量数据作为学习样本并自动从数据集中学习相应特征,对洗浴环境中的水汽、光照和目标的感知均表现出较好的鲁棒性。本研究团队之前使用了基于DCNNs的目标检测方法进行多部位皮肤检测研究。如:文献^[13]使用Faster R-CNN、YOLOv3、YOLOv4和CenterNet等算法进行皮肤检测,发现YOLOv4具有最高的平均精度均值(mean average precision, mAP),其值为78%;文献^[14]基于注意力机制改进了YOLOv4-tiny,并使用MobileNetV1、MobileNetV2

和 MobileNetV3 等算法对 YOLOv4 的主干网络进行轻量化处理,但所有改进模型的 mAP 值均低于 70%。因此,上述研究中的 mAP 值都不足以应对实际应用的需求。在这些研究基础上,本文使用其他目标检测算法进行多部位皮肤检测研究,扩充原始数据集,构建综合评价指标体系,并对模型性能进行评价,最后部署综合性能更好的模型在洗浴机器人中进行应用。

1 基于 DCNNs 的目标检测算法

目标检测是许多复杂计算机视觉工作的基础,具有非常广泛的应用场景。传统目标检测算法的精度取决于手工提取的特征,而手工提取的特征通用性差,检测精度难以提升,不适用于洗浴场景。随着大数据的爆发式增长和人工智能、高性能计算等领域计算能力的飞速提升,基于 DCNNs 的目标检测算法近年来成为主流,其性能远远优于传统算法^[15-16]。该算法提取的特征通用性好、表达能力强,对水汽、光照和目标的感知表现出更好的鲁棒性,更能满足洗浴场景的需求。根据两种分类标准,基于 DCNNs 的目标检测算法可分为两阶段/单阶段算法和基于锚框/无锚框算法。两阶段算法包括提取感兴趣区域特征阶段和基于此特征进行分类和回归阶段,经历两次边界框调整,算法精度高、速度慢。单阶段算法去除了候选框生成步骤,直接进行分类回归,调整一次边界框,算法精度低、速度快。基于锚框的算法在图像或者特征图上平铺大量预先定义好的先验框,对其坐标进行多次精炼得到边界框。为了改善锚框算法对不同研究集适应性差的问题,有研究人员提出了无锚框算法^[17]。如表 1 所示, YOLOv1、YOLOv6 和 YOLOX 是无锚框算法,缺少边界框的先验信息,对场景的泛化性更好。YOLOv2、YOLOv3、YOLOv4、YOLOv5 和 YOLOv7 是基于锚框的算法,使用固定大小的锚框作为边界的先验框。

单阶段目标检测算法具有良好的实时性,在实际应用场景中获得了普遍关注。在之前工作^[13-14]基础上,本研究选择表 2 中的算法执行多部位皮肤检测。YOLOv4 在 3 个尺度上执行预测,主干网络为 CSPDarknet53,加入 SPP 和 PANet 结构,通过连续上采样、下采样和通道叠加实现有效的

表 1 基于 DCNNs 的目标检测算法

Tab.1 Object detection algorithms based on DCNNs

算法	两阶段	单阶段
基于锚框	Faster R-CNN, R-FCN	SSD, YOLOv2/v3/v4/v5/v7, RetinaNet, RefineDet
无锚框	R-CNN, SPPNet, Fast R-CNN, DeNet, CPNDet	YOLOv1/v6/X, Cornernet, CenterNet, ExtremeNet, RepPoints, CentripetalNet, DenseBox, FCOS, FSAF, FoveaBox

表 2 本文选用的目标检测算法

Tab.2 Object detection algorithms selected by this article

算法类型	主干网络
YOLOv4	CSPDarknet53
YOLOv5s	New CSPDarknet53
YOLOX	CSPDarknet
YOLOv7	New ELANCSPP

特征融合。该算法采用 Mosaic 数据增强、自对抗训练和 DropBlock 正则化技术,以及标签平滑、学习率余弦退火衰减等方法,回归损失运用 CIOU 损失函数^[18],以避免 IoU 和 GIOU 损失导致的收敛速度慢和定位精度低等问题。YOLOv5 有多个版本,考虑到嵌入式硬件部署的便利性,本文选择 YOLOv5s 5.0 版进行研究。该版本算法主干网络为 CSP 结构,基于 SPP 和 PAN 结构加强特征融合,通过 Focus 操作降低计算量并保持较高精度,并采用 CIOU 损失函数,引入根据训练数据自动学习的自适应锚框策略。另外,运用跨网格匹配的正负样本分配策略,将正样本分配到多个网格中,提高网络训练的效率,保证正负样本的平衡^[19]。YOLOX 的主干网络使用 SiLU 激活函数、Focus 结构、SPP 结构和 CSP 结构,基于 FPN 结构加强特征提取。YOLOX 分开处理分类和回归任务,并引入 SimOTA 正负样本分配策略,动态分配正样本,自动分析每个 GT(ground truth)拥有的正样本数量,进一步提高检测精度^[20]。YOLOv7 的主干网络通过 ELAN 结构和由常规卷积与 Maxpool 双路径构成的 MPCConv 结构,结合 SPPCSPC 模块的多尺度特征融合及 SiLU 激活函数的平滑门控机制,实现了高效特征提取与空间上下文建模。训练时使用 SimOTA 算法,自适应匹配真实框与先验框数量,有效平衡了检测精度与训练效率^[21]。

2 面向洗浴场景的皮肤检测数据集

收集包含皮肤的图像，并考虑位置、光照、模糊度和水汽等因素以增强数据集多样性。原始数据集包括 1000 个图像，它是以往研究^[13-14]所使用的数据集。图 1 为数据集中的示例图像。



图 1 数据集示例图像

Fig.1 Example images from the datasets

使用图像标注工具 Labellmg，手动绘制矩形框以指定目标的标签和边界，用以表示目标的类别和确切位置，为每个图像生成 PASCAL VOC 格式的相应 XML 文件。XML 文件包含有关图像的文件名、GT Box 的坐标和标签等信息，Face_skin(脸)、Trunk_skin(躯干)、Upperlimb_skin(上肢)、Lowerlimb_skin(下肢)、Hand_skin(手)、Foot_skin(足) 表征人体不同部位的皮肤区域。研究发现，原始数据集存在类别不平衡现象，为此，基于离线数据增强方法对原始数据集进行扩充，使各类别数量比例接近 1，构成新数据集。图像数由 1000 个增至 2266 个，各类别数量如表 3

表 3 原始数据集和新数据集中各皮肤区域的数量

Tab.3 The number of each skin region in the original datasets and the new datasets

皮肤区域	原始数据集	新数据集
Face_skin	972	972
Trunk_skin	144	932
Upperlimb_skin	1060	1092
Lowerlimb_skin	525	997
Hand_skin	1161	1161
Foot_skin	200	905

所示。

对 2266 个图像进行统计发现，肤色白皙、中等、偏暗的图像数依次为 745、791、730 个，光照充足和不充足的图像数依次为 1209、1057 个，而有水汽和无水汽的图像数依次为 1064、1202 个。

3 模型训练及结果分析

3.1 迁移学习

从头开始训练网络需要大量标注数据^[22]，但手工标注费时费力、效率低、错误率高。将小数据集与迁移学习相结合可以快速训练出可接受的模型^[23-24]，达到从头开始训练网络的等效性能，减小模型过拟合的风险^[25]。本研究使用 ImageNet 数据集对所有目标检测模型进行预训练，将训练得到的权重文件作为初始权重。

3.2 训练环境及参数

使用 Pytorch 框架在上海理工大学超级计算机中心训练所有模型。训练集、验证集和测试集的比例分别为 60%、20%、20%。学习率初始值设置为 0.001，衰减率设置为 0.01，使用随机梯度下降 (stochastic gradient descent, SGD) 作为网络训练的优化器。Batch size 为 64，表示每次输入到网络中进行训练的图像数量。在自建的目标检测数据集上训练时，首先冻结预训练模型中主干网络的参数，训练其他参数，然后释放对主干网络参数的限制并训练所有参数，以便对参数进行微调。这种训练策略提高了网络的收敛速度和训练效率。

3.3 综合评价指标

召回率 R (recall) 和精度 P (precision) 是评价网络性能的指标，如式 (1)、(2) 所示^[26]。以召回率为横轴、精度为纵轴，绘制 P - R 曲线，其与坐标轴围成的面积为平均精度 (average precision, AP) 值，代表精度和召回率的平衡指标。mAP 值是所有类别 AP 的均值，是多类别目标检测的重要评价指标之一。在洗浴机器人中应用时，模型大小影响实际部署性能，而权重文件一定程度上可以反映模型的大小，所以选择 mAP 值和权重大小作为评价模型性能的基础指标。

$$R = T_p / (T_p + F_N) \quad (1)$$

$$P = T_p / (T_p + F_p) \quad (2)$$

式中， T_p 、 F_p 、 F_N 分别指模型预测为正类的正样本数量、模型预测为正类的负样本数量、模型预

测为负类的正样本数量。

建立综合评价指标 W , 评估精度和模型大小的平衡性, 以单位权重可以产生多少精度为切入点, 计算方法如式(3)所示。

$$W = \frac{B'}{A'} \quad (3)$$

式中: A' 表示权重大小; B' 表示模型的 mAP 值。 W 越大表征模型综合性能越好。

3.4 结果分析

模型训练结束后, 使用最优模型权重文件对测试集进行皮肤检测, 结果如表4所示。由各类别的 AP 数据可知, YOLOv7 达到了最佳的脸、足、手、下肢和上肢的检测效果, YOLOv5s 和 YOLOv7 实现了躯干最佳检测效果。YOLOv4、YOLOv5s、YOLOX、YOLOv7 的 mAP 值均在 85% 以上, 其中, YOLOv5s 和 YOLOv7 的 mAP 值达到了 90% 及以上。但 YOLOv7 的权重文件大于 140 MB, YOLOv4 的权重文件大于 240 MB。为了评估精度和模型大小的平衡性, 计算综合评价指标结果如表5所示。表中数据显示, YOLOv5s 的 W 值最大, 在 mAP 值和模型大小之间达到了最好的平衡。

表4 基于新数据集训练的算法性能

Tab.4 Algorithm performance based on training on the new datasets

算法类型	AP值/%						mAP值/%	权重文件大小/MB
	脸	足	手	下肢	躯干	上肢		
YOLOv4	95	87	86	84	92	70	86	244
YOLOv5s	97	94	89	90	93	76	90	27
YOLOX	96	91	88	90	92	70	88	35
YOLOv7	98	96	93	91	93	81	92	143

表5 综合评价指标结果

Tab.5 Results of comprehensive evaluation index

算法类型	A'	B'	W
YOLOv4	244	86	0.35
YOLOv5s	27	90	3.33
YOLOX	35	88	2.51
YOLOv7	143	92	0.64

为了评估数据集扩充对模型性能的影响, 基

于原始数据集训练了表2中的所有算法类型, 并与基于新数据集训练的结果进行对比, 如表6所示。表中数据显示, 在网络权重文件大小不变的情况下, YOLOv4、YOLOv5s、YOLOX 和 YOLOv7 的 mAP 值分别从 78%、69%、67%、71% 提升至 86%、90%、88%、92%, 精度平均提升 18%。由此可见, 原始数据集的类别不平衡问题解决之后, 模型大小不变, 而性能却能得到明显提升。

表6 基于原始和新数据集训练的算法性能对比

Tab.6 Comparison of algorithm performance based on the original and new datasets

算法类型	数据集	mAP值/%	权重文件大小/MB
YOLOv4	原始数据集	78	244
	新数据集	86	244
YOLOv5s	原始数据集	69	27
	新数据集	90	27
YOLOX	原始数据集	67	35
	新数据集	88	35
YOLOv7	原始数据集	71	143
	新数据集	92	143

4 算法部署及性能测试

采用人体模型代替真实人体开展实验。在 Tesla T4 和 TX2 平台上部署了综合性能最好的 YOLOv5s 进行检测。该算法处理大小为 640×480 的一帧 RGB 图像所需时间为 14 ms, 帧率约为 71 fps。在 TX2 平台上, 将 YOLOv5s 的 pt 权重文件转化为硬编码形式、后缀名为 wts 的权重文件, wts 文件再转化为推理引擎。调用推理引擎对 192 个分辨率为 640×480 的 RGB 图像进行检测, 发现 YOLOv5s 单个图像平均检测时间为 23.7 ms, 即帧率为 42 fps。由帧率可知, YOLOv5s 部署在 Tesla T4 和 TX2 平台上均能达到实时检测。

实际洗浴场景中充满水汽, 这种环境的模拟在实验室中具有一定的难度。因此, 本文采用了图像处理算法对包含人体模型的图像进行处理, 以生成具有不同水汽效果的场景图, 从而模拟实际洗浴场景中的水汽环境。为了验证该方法的有效性, 使用 YOLOv5s 对不同水汽效果的场景图

进行了检测,结果如图2所示。在含水汽的环境中, YOLOv5s 仍然能够识别 Trunk_skin 和 Upperlimb_skin 类别, 对应于人体的背部和上肢部位。值得注意的是, 由于实验的需要, 图2中的人体模型已被部分去除, 导致图2的检测结果未能涵盖所有6个类别。深度学习技术具有较强的鲁棒性, YOLOv5s 在水汽环境中仍能准确地检测到人体不同部位的皮肤。

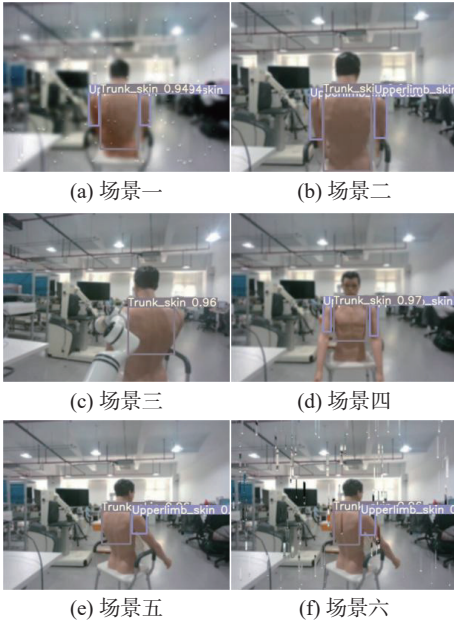


图2 水汽场景图的检测结果

Fig.2 Detection results of the water vapor scene images

5 洗浴机器人联合实验

5.1 视觉系统

考虑到性价比, 最终选择 TX2 平台作为部署设备。将视觉传感器 D455 通过 USB 连接到 TX2 平台上, 在 TX2 平台上借助深度学习加速技术栈 CUDA、cuDNN 和 TensorRT 等实现对 YOLOv5s 的处理并输出皮肤检测结果, 采用网口实现视觉系统和洗浴机器人的通信。为了防止水汽对相机的侵蚀, 采用眼在手外的方式, 将 D455 固定在可观察到视野中目标的位置处, 而不放置在洗浴机器人末端。视觉系统配置及技术指标如表7所示。

5.2 目标点三维位姿建模

由相机投影模型可知, 式(4)成立。其中: Z_c 为目标点的距离; $[u, v]$ 为目标点的像素坐标; $[X_c, Y_c, Z_c]$ 为目标点在 RGB 相机坐标系中的三维位置; K_{int} 为相机内参矩阵^[27]。如果已知

表7 视觉系统技术指标

Tab.7 Technical specifications of the vision system

配置	技术指标
主处理器	NVIDIA TX2
连接相机数量	1路
RGB图像分辨率	640×480
感知目标	助浴场景下的用户
通信接口	以太网
存储	联想500G SSD固态硬盘SATA3.0口

Z_c 和 K_{int} , 则可以由 RGB 图像中的 $[u, v]$ 获知其三维位置。可见, 从 RGB 图像中得到目标点的二维像素坐标, 并从 D455 输出的深度图中得到目标点处的深度值, 即可完成对目标点在 RGB 相机坐标系下三维位置 $[X_c, Y_c, Z_c]$ 的建模。目标点处的姿态则需要考虑目标点周围的信息。如图3所示, O_0-uv 是相机模型中的像素坐标系, 假设绿色为洗浴机器人执行定位的目标点, 以目标点为圆心, 以一定尺寸为半径绘制圆。红色为圆内所有与目标点邻近的像素点, 则结合相应深度值即可计算得到目标点和邻近点在 RGB 相机坐标系下的三维位置。它们将组成空间中身体表面的曲面, 通过此三维点集的法线即可确定目标点姿态。

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_{int} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (4)$$

由于视觉系统输出的目标点三维位姿位于 RGB 相机坐标系下, 为了给洗浴机器人控制系统提供必要的信息, 需要将目标点位姿转换至机器人基坐标系下, 因此, 进行了手眼标定^[17]。标定系统包括洗浴机器人、D455 和一块基于开源库 ArUco library 的 ArUco 标定板。相机固定在安装底座上, ArUco 标定板固定在机械臂末端, 通过调节机械臂位姿使标定板处于相机视野内。令机器人末端依次到达 17 个不同位姿, 并识别相机视野内的 ArUco 标记, 利用功能包 easy_handeye 求解得到相机在机器人基坐标系下的位姿。标定结果如表8所示, 保留至小数点后3位。

姿态四元数可以转换为旋转矩阵, 如式(5)所示。 r 为基坐标系在 RGB 相机坐标系下的旋转矩阵。RGB 相机坐标系相对基坐标系的转换矩阵 M 可表示为式(6), R 计算方法如式(7)所示。

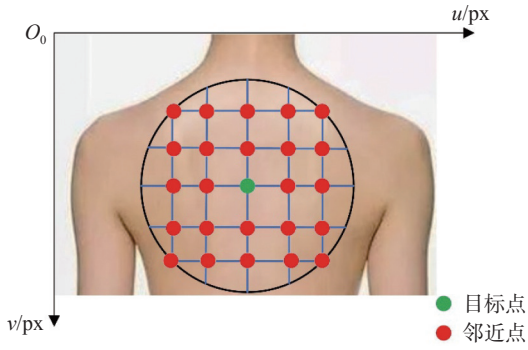


图3 基于邻近点信息的姿态确定示意图

Fig.3 Schematic diagram of attitude determination based on proximity point information

表8 手眼标定结果

Tab.8 Results of hand-eye calibration

输出参数	分量	数值
平移向量	t_x	0.250
	t_y	0.015
	t_z	-0.031
姿态四元数	x	-0.068
	y	-0.045
	z	0.703
	w	0.707

$$\mathbf{r} = \begin{bmatrix} 1 - 2y^2 - 2z^2 & 2(xy + wz) & 2(xz - wy) \\ 2(xy - wz) & 1 - 2x^2 - 2z^2 & 2(xz - wy) \\ 2(xz + wy) & 2(yz - wx) & 1 - 2x^2 - 2y^2 \end{bmatrix} \quad (5)$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \quad (6)$$

$$\mathbf{R} = \mathbf{r}^T \quad (7)$$

式中, \mathbf{t} 为平移向量。

5.3 联合实验及实验结果

由于背部区域是人类洗浴时较难擦洗到的部位, 本文主要开展了针对背部洗浴任务的实验。实验平台由视觉系统、人体模型、洗浴椅、无线网卡和洗浴机器人组成, 如图4所示。YOLOv5s将多部位皮肤检测结果发布为话题 boundingboxes_tensor, 包括皮肤区域的宽 λ_w 、高 λ_h , 以及左上角点坐标 λ_1 、 λ_t , 如图5所示。目标点是皮肤区域的中心 O_m , O_m 处的像素坐标如式(8)所示。根据目标点三维位姿建模方法可得到目标点在机器人坐标系下的三维位姿, 由机器人控制器控制其运行至目标点 O_m 。

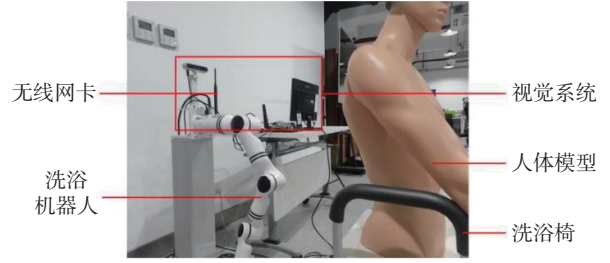


图4 实验平台

Fig.4 Experimental platform

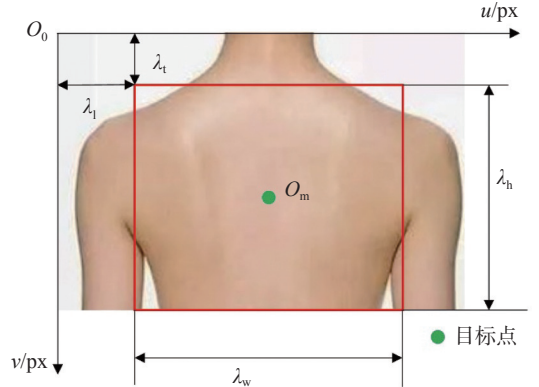


图5 boundingboxes_tensor 映射到背部区域目标点的示意图
Fig.5 Schematic diagram of boundingboxes_tensor mapping to the target points in the back region

$$\begin{cases} u_{O_m} = \lambda_1 + \lambda_w/2 \\ v_{O_m} = \lambda_t + \lambda_h/2 \end{cases} \quad (8)$$

改变人体模型位置, 重复进行目标点定位实验 50 次。结果显示, 成功次数为 46 次, 成功率为 92%。其中, 4 次失败中, 2 次因背部区域被机器人遮挡导致目标点检测错误, 2 次因人体模型在机器人运动期间出现意外移动而导致任务失败。

为了解决上述问题, 可将能够同时产生 RGB 和点云的 RGB-D 相机应用在机器人中, 将点云作为监督信息, 反映背部区域被机器人遮挡、人体模型在机器人运动期间出现意外移动时的情况。上述情况下, 点云三维点坐标信息将发生较大波动, 检测到点云三维坐标发生较大幅度波动, 即发出预警, 停止此次定位, 并将其作为开关条件, 启动新一轮的多部位皮肤检测, 重新寻找目标点进行定位。在使用点云作为监督信息后, 定位成功率达到了 100%, 有效消除了背部区域被机器人遮挡和人体模型意外移动造成的不利影响。

6 结束语

经扩充数据集后, 使用原始数据集和新数据集对多部位皮肤检测进行了研究, 并建立了以

mAP 值和模型大小为基础指标的综合评价指标。结果显示, YOLOv5s 在 mAP 值和模型大小之间达到了更好的平衡, 将其部署在 Tesla T4 和 TX2 平台上均能实现实时检测。考虑到性价比, 基于 TX2 平台集成了视觉系统和洗浴机器人, 并完成了目标点三维位姿建模。基于目标检测对洗浴任务中的皮肤进行分部位检测, 并控制机器人运行至背部区域, 任务成功率为 92%, 使用点云作为监督信息可将成功率提升至 100%。未来可以在洗浴机器人末端增设力传感器, 在机器人抵达目标皮肤区域后, 根据力传感器的反馈精细地调整机器人的运动, 以实现全自动、全覆盖洗浴模式。

参考文献:

- [1] 冯士凯, 李平, 许朋, 等. 洗浴机器人擦洗轨迹规划与实时控制 [J]. 建模与仿真, 2024, 13(1): 123–135.
- [2] HASSAN E K, SAUD J H. HSV color model and logical filter for human skin detection[J]. AIP Conference Proceedings, 2023, 2457(1): 040003.
- [3] NAJI S, JALAB H A, KAREEM S A. A survey on skin detection in colored images[J]. *Artificial Intelligence Review*, 2019, 52(2): 1041–1087.
- [4] GANESAN P, SATHISH B S, LEO JOSEPH L M I, et al. HSV model based skin color segmentation using uncomplicated threshold and logical and operation[C]//Proceedings of the 9th International Conference on Advanced Computing and Communication Systems (ICACCS). Coimbatore: IEEE, 2023: 415–419.
- [5] FOTOUHI M, ROHBAN M H, KASAEI S, et al. Skin detection using contourlet-based texture analysis[C]//Proceedings of the 4th International Conference on Digital Telecommunications. Colmar: IEEE, 2009: 59–64.
- [6] GKALTEMIS M A, LYMPERAKIS G, GENNIS G, et al. A hardware-friendly low-power area-efficient GMM-based analog classifier for skin detection[C]//Proceedings of 2022 Panhellenic Conference on Electronics & Telecommunications (PACET). Tripolis: IEEE, 2022: 1–4.
- [7] KAWULOK M, KAWULOK J, NALEPA J. Spatial-based skin detection using discriminative skin-presence features[J]. *Pattern Recognition Letters*, 2014, 41: 3–13.
- [8] ROY K, SAHAY R R. A robust multi-scale deep learning approach for unconstrained hand detection aided by skin segmentation[J]. *The Visual Computer*, 2022, 38(8): 2801–2825.
- [9] KARIM A H M Z, MIAH M S, JAMAL G R A, et al. Application of feature based face detection in adaptive skin pixel identification using signal processing techniques[C]//Proceedings of the 24th International Conference on Computer and Information Technology (ICCI). Dhaka: IEEE, 2021: 1–5.
- [10] KHAN M A, MUHAMMAD K, SHARIF M, et al. Multi-class skin lesion detection and classification via teledermatology[J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(12): 4267–4275.
- [11] KHAN M A, SHARIF M, AKRAM T, et al. Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization[J]. *Diagnostics*, 2021, 11(5): 811.
- [12] ZLATINTSI A, DOMETIOS A C, KARDARIS N, et al. I-Support: a robotic platform of an assistive bathing robot for the elderly population[J]. *Robotics and Autonomous Systems*, 2020, 126: 103451.
- [13] LI P, YU H L, LI S J, et al. Comparative study of human skin detection using object detection based on transfer learning[J]. *Applied Artificial Intelligence*, 2021, 35(15): 2370–2388.
- [14] LI P, HAN T Y, REN Y F, et al. Improved YOLOv4-tiny based on attention mechanism for skin detection[J]. *PeerJ Computer Science*, 2023, 9: e1288.
- [15] RASHID M, KHAN M A, ALHAISONI M, et al. A sustainable deep learning framework for object recognition using multi-layers deep features fusion and selection[J]. *Sustainability*, 2020, 12(12): 5037.
- [16] RASHID M, KHAN M A, SHARIF M, et al. Object detection and classification: a joint selection and fusion strategy of deep convolutional neural network and SIFT point features[J]. *Multimedia Tools and Applications*, 2019, 78(12): 15751–15777.
- [17] SUN Y W, PAN B, FU Y L. Development of a novel hand-eye calibration for intuitive control of minimally invasive surgical robot[J]. *Frontiers of Mechanical Engineering*, 2022, 17(3): 42.
- [18] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 2020: 12993–13000.
- [19] 杨明轩, 陈琳. 改进 YOLOv5 的棉田杂草检测 [J]. 现代电子技术, 2024, 47(24): 60–67.
- [20] 吕志轩, 魏霞, 马志钢. 改进 YOLOX 的轻量级安全帽检测方法 [J]. *计算机工程与应用*, 2023, 59(1): 61–71.
- [21] 张永智, 何可人, 戈珏. 改进 YOLOv7 网络在低空遥感图像目标检测中的应用 [J]. *计算机工程与科学*, 2024, 46(7): 1269–1277.

- contraction in humans as revealed by magnetoencephalography[J]. *The Journal of Physiology*, 2010, 527(3): 623–631.
- [3] BECK T W, HOUSH T J, JOHNSON G O, et al. Mechanomyographic and electromyographic responses during submaximal to maximal eccentric isokinetic muscle actions of the biceps brachii[J]. *Journal of Strength and Conditioning Research*, 2006, 20(1): 184–191.
- [4] BARATTA R V, SOLOMONOW M, ZHOU B H, et al. Methods to reduce the variability of EMG power spectrum estimates[J]. *Journal of Electromyography and Kinesiology*, 1998, 8(5): 279–285.
- [5] 王秀汝, 赵文汝, 刘金敬, 等. 肌电生物反馈治疗中枢神经系统损伤所致腕、踝关节异常运动模式的临床研究 [J]. *中国康复医学杂志*, 2004, 19(2): 108–110.
- [6] 甄鑫, 董小雷, 蔡玉强. 人机环境中下肢关节扭矩动态估计的研究 [J]. *科技与创新*, 2023, (13): 26–29.
- [7] 史小华. 坐/卧式下肢康复机器人研究 [D]. 秦皇岛: 燕山大学, 2014.
- [8] 任小梅, 杨刚. 基于模糊聚类技术的肌电信号完全分解算法 [J]. *计算机应用*, 2016, 36(3): 878–882.
- [9] 王金玮, 曹乐, 阚秀, 等. 基于表面肌电信号的 LDA-BPNN 双臂手势识别算法 [J]. *传感器与微系统*, 2023, 42(6): 158–160, 168.
- [10] 孔康, 李德盈, 孙中圣. 基于小波包组合特征和 LMS-LSTM 的表面肌电信号分类 [J]. *电子技术应用*, 2022, 48(10): 92–96.
- [11] 陆伟. 基于肌电信号的人体上肢肌力预测方法研究 [D]. 合肥: 中国科学技术大学, 2023.
- [12] NAEEM U J, XIONG C H, ABDULLAH A A. EMG-muscle force estimation model based on back-propagation neural network[C]//2012 IEEE International Conference on Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS) Proceedings. Tianjin: IEEE, 2012: 222–227.

(编辑: 丁红艺)

(上接第 9 页)

- [22] PATTNAIK G, SHRIVASTAVA V K, PARVATHI K. Transfer learning-based framework for classification of pest in tomato plants[J]. *Applied Artificial Intelligence*, 2020, 34(13): 981–993.
- [23] KHAN M A, ZHANG Y D, SHARIF M, et al. Pixels to classes: intelligent learning framework for multiclass skin lesion localization and classification[J]. *Computers & Electrical Engineering*, 2021, 90: 106956.
- [24] CHATO L, REGENTOVA E. Survey of transfer learning approaches in the machine learning of digital health sensing data[J]. *Journal of Personalized Medicine*, 2023, 13(12): 1703.
- [25] KHAN M A, AKRAM T, ZHANG Y D, et al. Attributes based skin lesion detection and recognition: a mask RCNN and transfer learning-based deep learning framework[J]. *Pattern Recognition Letters*, 2021, 143: 58–66.
- [26] 仇旭阳, 黄影平, 郭志阳, 等. 基于深度学习的障碍物检测与深度估计 [J]. *上海理工大学学报*, 2020, 42(6): 558–565.
- [27] SHI Z C, SHANG Y, ZHANG X F, et al. DLT-lines based camera calibration with lens radial and tangential distortion[J]. *Experimental Mechanics*, 2021, 61(8): 1237–1247.

(编辑: 丁红艺)