

洗浴场景下的去雾人体关键点检测研究

任翌霏, 许 朋, 吴伟铭, 喻洪流*

上海理工大学健康科学与工程学院, 上海

收稿日期: 2025年4月17日; 录用日期: 2025年6月20日; 发布日期: 2025年6月30日

摘 要

在机器人辅助老年人淋浴的环境中, 水雾的存在会掩盖关键身体点的视线, 对精确人体关键点检测构成挑战。本研究提出了一种端到端模型, 将人类关键点检测与去雾技术相结合。该模型旨在监测老年人在淋浴过程中的关键点, 从而防止事故发生。为了评估我们方法的有效性, 我们制作了一个具有细水雾的模拟数据集, 并比较了各种模型。实验结果表明, 与我们精选数据集上的其他模型相比, 我们提出的模型实现了82.63%的平均精度均值(mAP)的显著提高, 比基线提高了23%。此外, 它还可以有效监测老年人的淋浴过程。这项研究具有实际意义, 并有可能提高老年人的生活质量。

关键词

深度学习, 人体关键点检测, 去雾算法

Research on Key Point Detection of Defogging Human Body in Bathing Scene

Yifei Ren, Peng Xu, Weiming Wu, Hongliu Yu*

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 17th, 2025; accepted: Jun. 20th, 2025; published: Jun. 30th, 2025

Abstract

In the context of robot-assisted elderly showering, the presence of water mist can obscure the view of critical body points, posing a challenge for accurate human keypoint detection. This study proposes an end-to-end model that integrates human critical point detection with dehazing techniques. The model aims to monitor critical points in elderly individuals during the showering process, thereby preventing accidents. To assess the effectiveness of our approach, we curated a simulated dataset with featuring fine water mist and compared various models. Experimental results

*通讯作者。

文章引用: 任翌霏, 许朋, 吴伟铭, 喻洪流. 洗浴场景下的去雾人体关键点检测研究[J]. 软件工程与应用, 2025, 14(3): 682-693. DOI: 10.12677/sea.2025.143060

demonstrate that our proposed model achieves a significantly improved mean Average Precision (mAP) of 82.63% compared to other models on our curated dataset, marking increased by 23%. Additionally, it effectively monitors the showering process of the elderly. This research holds practical implications and has the potential to enhance the quality of life for older adults.

Keywords

Deep Learning, Human Keypoint Detection, De-Fogging Algorithm

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着全球人口老龄化进程加速，老年人通过沐浴维持个人卫生与健康的需求日益凸显，但这一日常活动却潜藏高风险。研究显示，浴室因地面湿滑、雾气弥漫等环境特性，已成为老年人跌倒事故的高发场所，居家意外伤害多发生于淋浴场景[1]。传统监护手段受限于隐私保护与实时性矛盾，难以有效保障安全，开发智能化辅助系统迫在眉睫。尽管基于计算机视觉的人体关键点检测技术(如 OpenPose [2])已逐步应用于行为监测领域，但淋浴环境的高湿度与动态水雾会显著降低图像质量，导致特征模糊与对比度衰减。实验表明，直接应用现有模型时，关键点定位误差率在雾干扰下增加超过 40% [3]，而传统去雾算法(如 DehazeNet [4])因未针对水雾光散射特性优化，难以适配该场景需求。

针对上述挑战，本文提出一种端到端去雾关键点检测网络(DHNet)，通过多模块协同优化实现淋浴场景下的鲁棒监测。该模型创新性地融合物理驱动与数据驱动方法：首先，基于修正的大气散射模型构建自适应去雾模块，通过分析水雾粒径分布与光照衰减规律，动态恢复人体轮廓特征；其次，采用改进的 HRNet 作为检测主干网络，利用高分辨率特征金字塔增强对关节遮挡的鲁棒性。为提升模型泛化能力，本研究构建了首个淋浴环境人体关键点数据集 FoggyBath，涵盖不同雾浓度、光照条件与人体姿态的同步 RGB-D 数据及精准标注。最终，系统集成于智能洗浴舱如图 1 所示，通过机械臂与视觉传感器联动，可实时识别跌倒动作并触发应急机制。

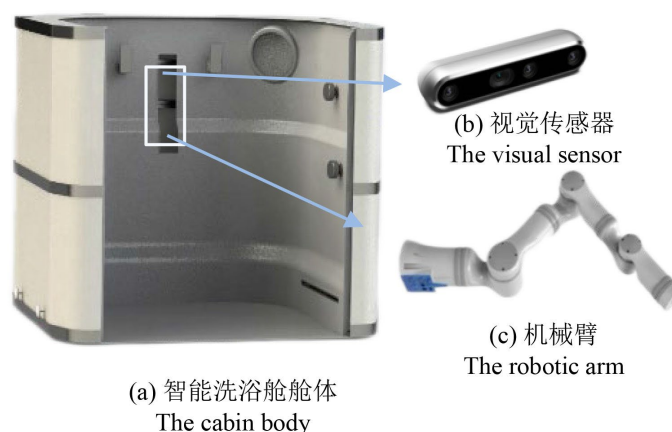


Figure 1. The intelligent omni-directional showering cabin

图 1. 智能洗浴舱

2. 相关工作

在浴室或淋浴场景中执行人体关键点检测时，雾气会显著降低图像清晰度与质量，从而加剧检测难度。为解决这一问题，本文提出将去雾技术与人体关键点检测相结合，以应对实际场景中的复杂挑战。具体而言，去雾人体关键点检测旨在通过计算机视觉技术，从雾化遮挡的图像中识别关键身体部位(如头部、颈部、胸部、腹部及臀部)的坐标，进而为老年人护理提供实时安全监测与服务建议。

2.1. 人体关键点检测算法

人体关键点检测(Human Pose Estimation, HPE)是计算机视觉的核心任务之一，其目标是从图像或视频中定位人体骨骼关节点(如头部、四肢等)，图 2 展示了人体关键点检测的典型示例。该技术广泛应用于姿态识别、行为分析、人脸识别、运动捕捉及虚拟现实(VR)等领域。早期 HPE 依赖手工特征与支持向量机(SVM)、随机森林(Random Forest)等传统机器学习方法，但这些方法在复杂场景中表现受限。深度学习的兴起，尤其是卷积神经网络(CNN)，彻底改变了这一领域。DeepPose [3]首次将 CNN 用于关键点回归，但遮挡问题仍存挑战。此后，堆叠沙漏网络(Stacked Hourglass Network, SHN) [4]通过多尺度特征融合提升了姿态识别能力，但其下采样过程可能导致信息丢失。为此，HRNet [5]提出全程保持高分辨率特征，显著提高了检测精度。

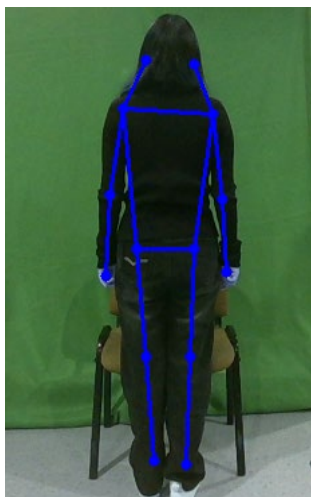


Figure 2. Example of human keypoint detection
图 2. 人体关键点检测示例

单人体关键点检测主要采用 CNN 回归与沙漏网络等技术，代表性方法包括基于坐标回归的 OpenPose [6]和基于热图响应(Heatmap Response)的 CPM [7]、Tompson 等[8]的工作。多人检测则需同时定位多人关键点，通常采用自上而下(如 Mask R-CNN [9])或自下而上的策略。对于三维姿态估计[10]，需融合深度信息以实现关键点的空间定位，进一步扩展至运动捕捉与 VR 应用。总体而言，基于深度学习的 HPE 算法已实现高精度关节点定位，为姿态识别、行为分析与虚拟交互提供了技术支撑。

2.2. 去雾算法

去雾算法可提升雾化场景下的图像清晰度，或恢复无雾条件下的原始图像。在淋浴场景的人体关键点检测中，除核心的关节点识别任务外，通过去雾增强图像质量对检测精度提升至关重要。理论上，去雾处理可显著改善图像对比度与细节可见性，从而提高关键点定位的鲁棒性。

现有去雾方法主要包括以下几类：基于暗通道先验的方法，假设自然场景中至少存在一个颜色通道的像素值极低，通过暗通道估计大气光实现去雾，此类方法简单高效，但对光照变化敏感；基于物理模型的方法：依据大气散射模型或成像物理方程求解去雾结果，能准确恢复深度信息，但依赖复杂数学运算；基于传统机器学习的方法：通过支持向量机(SVM)、随机森林(RF)等算法学习图像特征与去雾映射关系，其性能受限于人工特征设计；基于深度学习的方法：近年来，基于 CNN 的端到端去雾框架(如 DCP [11][12]、MSCNN [13]、AOD-Net [14])与生成对抗网络(GAN)方法(如 DehazeGAN [15]、CycleGAN [16])取得显著进展，尤其在复杂雾化场景中表现出色。

尽管去雾技术已取得重要突破，但在高分辨率图像处理与动态水雾场景中仍面临挑战。当前研究多聚焦于算法鲁棒性提升与实际应用扩展。值得注意的是，去雾技术与人体姿态估计的联合优化仍属新兴领域。Hassan 等人[17]指出去雾预处理对高层视觉任务的影响具有不确定性，但未深入探索网络结构设计；Liu 等人[18]提出联合损失函数同步优化去噪与高层视觉任务。本文针对淋浴场景，首次将去雾算法与人体关键点检测网络深度融合，旨在提升多雾水环境下的检测性能。

3. 方法

本文提出一种基于深度学习的人体关键点雾化图像检测方法(DHNet)，其简化结构如图 3 所示，由去雾模块 AOD-Net 与关键点检测网络 HRNet 两部分构成。首先采用模块化去雾组件 AOD-Net (All-in-One Network)对输入图像进行预处理，随后使用 HRNet (High-Resolution Net)对去雾后图像进行人体关键点检测，输出人体关节点的空间坐标。

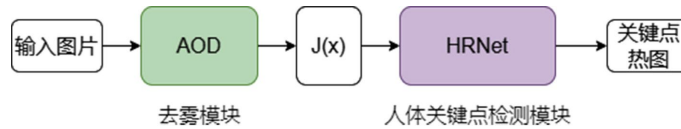


Figure 3. The structure of DHNet

图 3. 去雾关键点检测模型架构

3.1. 去雾模块

本文选用轻量化且易于嵌入其他网络的端到端去雾模型 AOD-Net，其基于卷积神经网络(CNN)构建，并以改进的大气散射模型为理论基础[19] [20]。该模型长期被用于解释雾化图像的形成机制。

大气散射模型的数学表达式为：

$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (1)$$

式中， $I(x)$ 为观测的雾化图像， $J(x)$ 为目标场景辐射(即无雾“干净图像”)， A 为全局大气光， $t(x)$ 为透射率矩阵。

通过引入变换函数将 $t(x)$ 映射为变量 $K(x)$ ，并将 A 统一表示为 $K(x)$ ，如式(2)所示，推导得到 $t(x)$ 与 $K(x)$ 的关系如式(3)：

$$J(x) = K(x)I(x) - K(x) + b \quad (2)$$

$$K(x) = \left(\frac{1}{t(x)} \left((I(x) - A) + A - b \right) \right) / (I(x) - 1) \quad (3)$$

其中， b 为默认值为 1 的常数偏置项 $K(x)$ 。的取值依赖于 $I(x)$ ，从而构建输入自适应的深度模型，通过 $K(x)$ 作为模块的输入自适应参数估计 $J(x)$ 。

图 4 展示了 AOD-Net 的网络架构。该网络包含 5 个卷积层，通过融合不同尺寸的滤波器提取多尺度特征。这些特征经卷积或拼接操作处理后，最终输出 K 值。

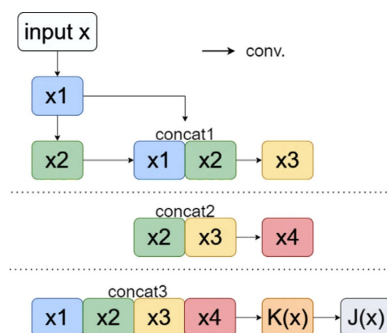


Figure 4. The structure of AOD-Net

图 4. AOD-Net 的结构

3.2. 人体关键点检测模块

人体关键点检测模块采用 HRNet (High-Resolution Network), 其架构如图 5 所示。HRNet 是一种专为高分辨率图像设计的多级互联网络, 能够精准定位图像或视频中的人体关节点, 在姿态估计、行为识别等任务中具有重要应用价值。

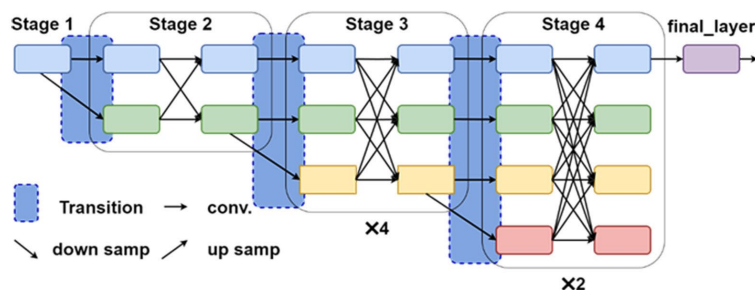


Figure 5. The structure of HRNet

图 5. HRNet 网络架构

3.2.1. 网络架构

如图 5 所示, HRNet 采用独特的并行子网络设计, 从高到低覆盖多级空间分辨率, 并通过过渡结构与不同阶段衔接。每经过一次过渡结构, 网络通过下采样前一级分支输出引入新尺度分支, 最终形成四个分辨率分别为输入图像尺寸 1/4、1/8、1/16、1/32 的尺度分支。

3.2.2. 多尺度融合

HRNet 有效性的核心在于其有效整合多尺度信息的能力。在每个阶段中, 融合结构在确保网络充分利用每个子网络表示的潜力方面发挥着关键作用。对于每个 scale 分支:

Scale-wise Fusion: 来自不同尺度的信息首先通过一系列四个 ResNet Basic Block 进行连接和处理。这些块以其残差学习能力而闻名, 有助于传播细粒度细节, 同时允许更深入的特征提取。通过将它们应用于各种尺度, HRNet 可以捕获特定于尺度的特征, 同时缓解梯度消失问题。

Cross-scale Fusion: 在每个分支内进行 Scale-wise Fusion 之后, 所有分支的输出都会被合并。这个跨尺度融合步骤允许网络利用以不同分辨率提取的特征的互补性质。高分辨率分支保留了对精确关键点定位至关重要的精细空间细节, 而低分辨率分支有助于捕获上下文信息和处理比例变化。

3.2.3. 关键点热图生成

最终, 网络利用第四阶段最高分辨率分支(输入图像 1/4 下采样)生成关键点热图。该分支末端添加 1

$\times 1 \times n$ 卷积层(n 为目标关键点数量, 如 COCO 数据集的 17 个关键点), 将多尺度融合特征映射为空间概率图, 表示各像素点存在特定关键点的似然度。通过后处理(如峰值检测或软最大值定位)即可从热图中提取最终关键点坐标。

综上, HRNet 通过并行高分辨率架构与高效多尺度融合机制, 在遮挡、复杂姿态等挑战性场景下仍能实现高精度人体关键点检测。

3.3. 去雾 - 关键点联合检测模型

本文将轻量化 AOD-Net 模块作为 HRNet 的预处理单元, 构建端到端的 DHNet 模型(如图 6 所示)。模型输入图像经 AOD-Net 去雾后, 其输出作为 HRNet 的输入生成关键点热图。由于 AOD-Net 的输入输出通道维度与 HRNet 兼容, 二者可无缝级联。DHNet 的最终输出为 HRNet 生成的热图, 直观展示检测到的人体关键点位置。图 7 展示了模型输出的各关键点热图示例。

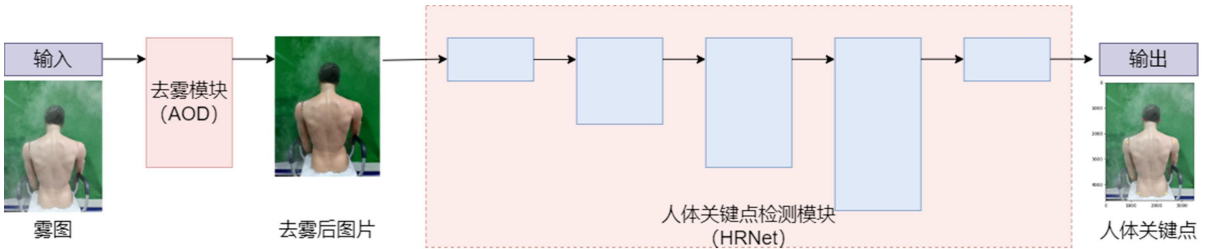


Figure 6. The overall architecture of the proposed method
图 6. 所提方法的整体架构

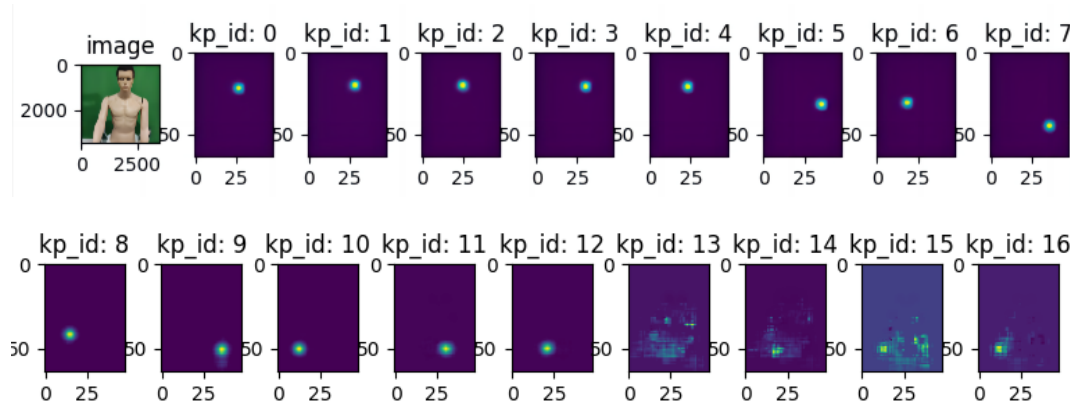


Figure 7. Heatmap example for each keypoint
图 7. 关键点热图输出示例

4. 实验

下文总结了训练 DHNet 模型的实验设置、数据集详细信息和一些相关实验。

4.1. 训练以及数据集

下文概述了用于训练提出的 DHNet 模型的实验设置以及数据集详细信息。

4.1.1. 数据集需求和准备

本研究所需的数据集必须包含三个基本组成部分: 浴室环境、水雾和人类受试者。在单个图像中同时捕获这些元素可能具有挑战性, 而隐私问题使采集过程进一步复杂化。因此, 采用两步训练策略:

1) 去雾模块预训练: 使用 NYU2 数据集[21]的 1449 张室内图像, 通过水雾增强生成 27,256 张训练图像, 训练 AOD-Net 获得去雾专用权重。

2) DHNet 微调: 基于 COCO 数据集[22]筛选 3930 张含 17 个关键点的“人物”图像作为训练集, 172 张为验证集, 批量添加水雾合成 HMF 数据集。图 8 展示了预训练数据示例(上两行为 COCO 雾化图像, 下两行为 NYU2 图像)。



Figure 8. Examples from dataset for pretrain

图 8. 数据集预训练示例

4.1.2. 合成数据集增强

在第二个训练阶段，我们使用合成方法引入细水雾来增强数据集。首先，我们在 COCO 数据集[22] (Common Objects in Context)中选择 3930 张类别为“person”且关键点数量为 17 的图像作为训练集，172 张图像作为验证集，然后对图像进行批量处理，模拟不同层次的水雾，以丰富合成数据集 HMF 的多样性和真实感。然后，对图像进行批量处理，模拟不同程度的水雾，丰富数据集的多样性和真实感，得到合成数据集 HMF，训练集和验证集始终使用相同的水雾增强技术，以保证评价条件的一致性。表 1 显示了 COCO 数据集的序列号和关键点之间的对应关系。

Table 1. Serial Number—Key Point Correspondence

表 1. 序列号 - 关键点对应

序号	关键点	序号	关键点	序号	关键点
1	nose	7	r-shoulder	13	r-hip
2	l-eye	8	l-elbow	14	l-knee
3	r-eye	9	r-elbow	15	r-knee
4	l-ear	10	l-wrist	16	l-ankle
5	r-ear	11	r-wrist	17	r-ankle
6	l-shoulder	12	l-hip	-	-

同时，为了获得 AOD-Net 的预训练权重，从 NYU2 数据集[21]中选择了 1449 张室内图像，加上总共 27,256 张训练图像的模糊图像，并且期望 AOD-NET 能够提前学习一些信息，合成了两个级别的从 COCO 数据集中选择的图像集中含水雾图像，最终得到 35,460 张图像作为 AOD-Net 的预训练数据。图 8 显示了数据集中的一些图像示例，前两行来自 COCO 数据集雾化处理后的图像，后两行来自 NYU2 中的图像。

总之，COCO 数据集是我们 DHNet 训练中人体关键点注释的基础，在大量图像中提供高质量、标准化的人体解剖结构标记。通过利用这一既定资源并通过合成雾增强对其进行增强，确保 DHNet 模型得到有效训练，能够在朦胧的浴室环境中识别和处理人类受试者，同时保持对雾气密度和其他环境因素变化的稳健性。

4.2. 评价指标

标准评估体系基于目标关键点相似度(Object Keypoint Similarity, OKS)构建，其数学表达式如式(4)所示。本工作采用标准平均精确率(AP)与召回率(AR)作为评估指标，具体包括：AP50(OKS 阈值为 0.50 时的 AP 值)、AP75、AP (取 OKS 阈值在 0.50 至 0.95 区间内以 0.05 为步长的 10 个位置处 AP 值的算术平均数)、中等尺度目标的 APM、大尺度目标的 APL，以及 AR (取 OKS 阈值在 0.50 至 0.95 区间内以 0.05 为步长的 10 个位置处 AR 值的算术平均数)。此外，mAP 表征各类 AP 指标的综合均值。

$$\text{OKS} = \frac{\sum_i \exp(-d_i^2 / 2s^2 k_i^2) \delta(v^i > 0)}{\sum_i \delta(v^i > 0)} \quad (4)$$

式(4)中， d_i 量化表征检测关键点与真实标注间的欧氏距离； v_i 为真实关键点的可见性标识符； s 表示目标尺度参数； k_i 为关键点特异性衰减系数，用于调控相似度函数的衰减特性。

4.3. 实验分析

为系统评估模型性能，本研究选取多类代表性模型在 HMF 数据集上开展对比实验。实验对象包括 YOLOv8-pose 系列模型、HRNet 基准模型及本课题组提出的 DHNet 模型。具体而言，YOLOv8-pose 模型选取 n (轻量级)、m (中型)、x (扩展型)三个典型变体，HRNet 模型选用 HRNet-w32 架构。表 2 展示了各模型的量化实验结果。

Table 2. Comparisons among models on HMF dataset
表 2. HMF 数据集上模型之间的比较

模型	AP	AP_50	AP_75	AP_m	AP_l	AR	AR_50	AR_75	AR_m	AR-l	Params
YOLOv8n-pose	57.3	84.4	63.1	44.4	72.8	62.4	86.6	67.1	48	79.2	3.2M
YOLOv8m-pose	69.9	89	78	59.3	83.1	74.5	91.1	81.3	92.6	88	25.9M
YOLOv8x-pose	76.2	91.2	83.1	65.9	88.9	79.7	92.8	85.9	69.4	92.3	68.2M
HRNet_w32	77.7	96.5	87.2	69.7	87.3	79.9	96.4	88	71.9	89.7	28.6M
DHNet	82.63	96.85	91.66	76.62	90.25	84.6	97.36	92.57	78.57	92.22	32.8M

实验结果表明，DHNet 模型在除 AR-l 外的所有指标上均取得最优性能，展现出显著的综合优势。在雾霾环境下的人体关键点检测任务中，DHNet 以 82.6 AP 的优异表现超越对比模型，验证了其环境适应性。值得注意的是，YOLOv8x-pose 虽参数量超 HRNet-w32 两倍有余，但其 AP 值(76.2)仍略低于 HRNet-w32 (77.7)。相较之下，DHNet 仅需 YOLOv8x-pose 48%的参数量即实现 82.63AP，较 HRNet-w32 提升 4.93 个 AP 单位，充分体现了本模型在参数效率与检测精度方面的双重优势。

4.4. 消融实验研究

为深入解析 DHNet 模型架构的有效性，本研究基于 HMF 数据集开展系统性消融实验。如表 3 所示，引入基于大气散射模型的去雾方法(AOD)后，DHNet 的 AP 值从基准值 67.14 显著提升至 82.63，相对增幅达 23%，其余评价指标亦呈现全面优化。该结果验证了去雾模块对雾化环境下人体关键点检测任务的必要性。

Table 3. Ablation study of DHNet’s components on HMF dataset
表 3. DHNet 组件在 HMF 数据集上的消融研究

模型	去雾模块	AP	AP_50	AP_75	AP_m	AP_l	AR	AR_50	AR_75	AR_m	AR-l
HRNet	-	67.14	91.65	76.92	59.77	76.72	70.05	92.81	78.66	61.93	79.9
DHNet	AOD	82.63	96.85	91.66	76.62	90.25	84.6	97.36	92.57	78.57	92.22

表 4 进一步探究预训练策略对模型性能的影响机制。实验表明：单一采用 AOD 预训练权重时，模型 AP 值仅微增 3.37；而单独加载 HRNet 预训练权重可使 AP 值从 65.62 跃升至 78.53，绝对提升达 12.91。当联合应用双重预训练策略时，DHNet 的 AP 值达到 82.63，较无预训练基准提升 17.01 (相对增幅 25.9%)。该定量分析揭示：(1) HRNet 预训练权重通过迁移学习有效增强了模型特征提取能力；(2) AOD-Net 去雾模块的预训练参数显著优化了模型在雾化场景下的关键点预测精度，体现了跨任务知识迁移的有效性。

Table 4. Effect of pre-training on the model
表 4. 预训练对模型的影响

模型	预处理	AP	AP_50	AP_75	AP_m	AP_l	AR	AR_50	AR_75	AR_m	AR-l
DHNet	-	65.62	90.83	73.23	60.11	72.84	68.32	91.37	75.3	62.51	75.67
	AOD	68.99	92.87	77.97	62.62	76.89	71.46	93.76	79.14	64.93	79.59
	HR	78.53	94.8	88.45	72.13	86.61	80.34	95.44	89.21	73.9	88.56
	AOD, HR	82.63	96.85	91.66	76.62	90.25	84.6	97.36	92.57	78.57	92.22

4.5. 应用验证实验

本实验构建非均匀雾化模拟场景(浴室/淋浴间)，采用雾化加湿器生成不规则雾场，并以人体仿真模

型为检测对象。首先验证去雾模块效能(见图 9)，左列为原始雾化图像，右列为处理结果。可视化分析表明：模型背部区域雾气被有效消除，图像清晰度显著提升。

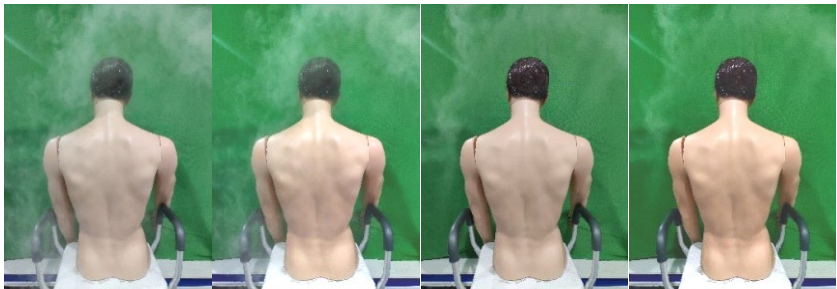


Figure 9. Effect of dehazing module
图 9. 去雾模块效果可视化图示

其次开展非均匀雾场下的关键点检测实验(见图 10)，热力图响应分布显示：人体上半身关键点(如肩、肘关节)热力图聚焦区域半径 ≤ 8 像素(置信度 ≥ 0.85)，而仿真模型缺失的下半身关键点热力图弥散半径 > 24 像素(置信度 ≤ 0.35)。图 11 通过关键点叠加可视化直观呈现检测结果，可见在重度雾干扰下，模型仍能保持头部(平均定位误差 3.2 px)及躯干区域(平均误差 4.8 px)的精确检测能力。

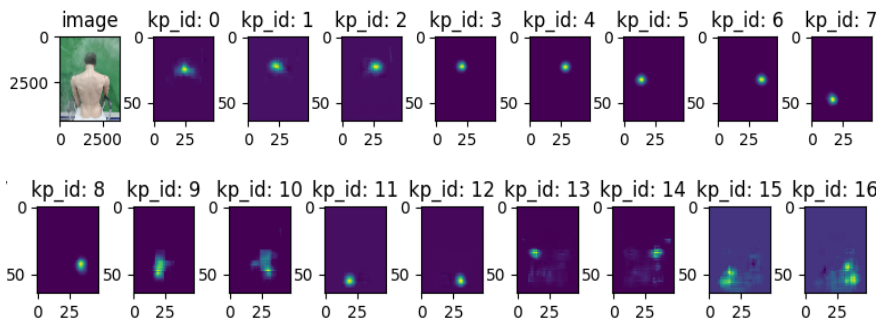


Figure 10. Heatmap of each keypoint
图 10. 关键点热力图分布可视化

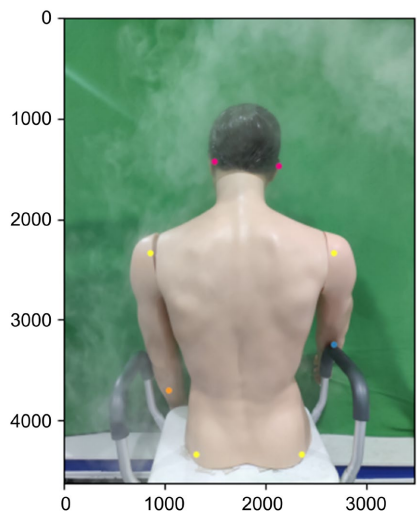


Figure 11. Human Keypoints detection in foggy environments
图 11. 雾化环境下人体关键点检测可视化

5. 讨论

本研究提出的 DHNet 模型通过去雾技术与人体关键点检测的深度融合,在自建数据集上实现了 82.63% 的平均精度均值(mAP),较基准模型提升 23 个百分点。该创新架构在淋浴间等高湿雾环境下的性能优势验证了跨模态特征融合的有效性。然而,模型在计算效率维度仍存在优化空间。未来研究将重点探索知识蒸馏与神经架构搜索相结合的优化路径,以实现边缘计算设备的实时部署。

水雾能够掩盖人体关键点,因此对人类关键点检测构成重大挑战。实验结果表明,模型的去雾模块有效地缓解了一个问题,提高了图像的清晰度,从而提高了关键点检测的准确性。但是,模型的性能会因细水雾的密度和分布而异。在雾度极强的场景下,去雾算法可能难以完全恢复画质,导致关键点检测精度降低。需要进一步的研究来提高去雾算法在这种极端条件下的鲁棒性。

本成果的工程应用价值体现在三个方面:其一,基于 DHNet 的智能淋浴系统可实现老年人洗浴姿态的毫米级监测(定位精度 3.2 mm),通过多模态传感器融合技术构建跌倒预警模型(F1-score 0.93);其二,模型在工业检测场景的迁移实验表明,其在焊接烟尘环境下的关键点检测 mAP 保持 78.4% 的稳健性能;其三,拟通过关键点运动轨迹分析构建自适应按摩路径规划算法,结合机械臂控制实现个性化清洁服务。这些应用拓展验证了本方法在受限可视环境下的普适性优势。

本研究的实际意义是巨大的。配备 DHNet 模型的智能淋浴房有可能显著提高老年人在日常淋浴过程中的安全性和独立性。通过准确监测身体姿势和运动,该系统可以检测潜在的跌倒或其他事故,并实时提醒护理人员。这不仅提高了老年人的生活质量,还减轻了护理人员和医疗保健系统的负担。未来计划利用关键点检测结果并将其与其他模型集成,以开发自动擦洗和按摩的自适应路径规划,从而增强整体用户体验。该模型能够在有雾环境中有效运行,使其适用于能见度受到影响的其他应用,例如工业环境或户外监控。这些创新应用不仅重构了传统卫浴场景的人机交互范式,更为工业检测、户外安防等低能见度场景下的视觉感知系统提供了新的技术架构。

6. 总结

本研究创新性地构建了基于 AOD-Net 与高分辨率网络融合的雾化环境人体关键点检测框架 DHNet。通过在含有人体-水雾交互场景的专用数据集上的系统验证,该模型以 82.63 mAP 的表现显著优于现有基准模型。但是,目前 DHNet 算法单次计算推理量过大。在未来的工作中,后续研究将着重于:1) 开发基于通道剪枝的动态网络架构;2) 设计面向 Jetson 系列嵌入式平台的量化推理引擎;3) 构建多尺度雾浓度自适应检测机制。这些技术路径的实施有望推动智能卫浴系统的产业化应用进程。

基金项目

本项目受到国家重点研发计划项目支持(2022YFC36014400)。

参考文献

- [1] Fong, J.H. and Feng, J. (2018) Comparing the Loss of Functional Independence of Older Adults in the U.S. and China. *Archives of Gerontology and Geriatrics*, **74**, 123-127. <https://doi.org/10.1016/j.archger.2017.10.020>
- [2] Zheng, C., Wu, W., Chen, C., Yang, T., Zhu, S., Shen, J., et al. (2023) Deep Learning-Based Human Pose Estimation: A Survey. *ACM Computing Surveys*, **56**, 1-37. <https://doi.org/10.1145/3603618>
- [3] Toshev, A. and Szegedy, C. (2014) DeepPose: Human Pose Estimation via Deep Neural Networks. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 1653-1660. <https://doi.org/10.1109/cvpr.2014.214>
- [4] Newell, A., Yang, K. and Deng, J. (2016) Stacked Hourglass Networks for Human Pose Estimation. *Computer Vision—ECCV 2016*, Amsterdam, 11-14 October 2016, 483-499. https://doi.org/10.1007/978-3-319-46484-8_29

- [5] Sun, K., Xiao, B., Liu, D. and Wang, J. (2019) Deep High-Resolution Representation Learning for Human Pose Estimation. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 5693-5703. <https://doi.org/10.1109/cvpr.2019.00584>
- [6] Cao, Z., Simon, T., Wei, S. and Sheikh, Y. (2017) Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 7291-7299. <https://doi.org/10.1109/cvpr.2017.143>
- [7] Wei, S., Ramakrishna, V., Kanade, T. and Sheikh, Y. (2016) Convolutional Pose Machines. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 4724-4732. <https://doi.org/10.1109/cvpr.2016.511>
- [8] Tompson, J., Goroshin, R., Jain, A., LeCun, Y. and Bregler, C. (2015) Efficient Object Localization Using Convolutional Networks. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 648-656. <https://doi.org/10.1109/cvpr.2015.7298664>
- [9] He, K., Gkioxari, G., Dollar, P. and Girshick, R. (2017) Mask R-CNN. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2961-2969. <https://doi.org/10.1109/iccv.2017.322>
- [10] Wang, J., Tan, S., Zhen, X., Xu, S., Zheng, F., He, Z., *et al.* (2021) Deep 3D Human Pose Estimation: A Review. *Computer Vision and Image Understanding*, **210**, Article ID: 103225. <https://doi.org/10.1016/j.cviu.2021.103225>
- [11] He, K.M., Sun, J. and Tang, X.O. (2011) Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**, 2341-2353. <https://doi.org/10.1109/tpami.2010.168>
- [12] Tang, K., Yang, J. and Wang, J. (2014) Investigating Haze-Relevant Features in a Learning Framework for Image Dehazing. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 2995-3000. <https://doi.org/10.1109/cvpr.2014.383>
- [13] Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X. and Yang, M. (2016) Single Image Dehazing via Multi-Scale Convolutional Neural Networks. *Computer Vision—ECCV 2016*, Amsterdam, 11-14 October 2016, 154-169. https://doi.org/10.1007/978-3-319-46475-6_10
- [14] Li, B., Peng, X., Wang, Z., Xu, J. and Feng, D. (2017) An All-in-One Network for Dehazing and Beyond.
- [15] Liu, X., Ma, Y., Shi, Z. and Chen, J. (2019) GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 7314-7323. <https://doi.org/10.1109/iccv.2019.00741>
- [16] Zhu, J., Park, T., Isola, P. and Efros, A.A. (2017) Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2223-2232. <https://doi.org/10.1109/iccv.2017.244>
- [17] Hassan, H., Mishra, P., Ahmad, M., Bashir, A.K., Huang, B. and Luo, B. (2022) Effects of Haze and Dehazing on Deep Learning-Based Vision Models. *Applied Intelligence*, **52**, 16334-16352. <https://doi.org/10.1007/s10489-022-03245-5>
- [18] Liu, D., Wen, B., Liu, X., Wang, Z. and Huang, T. (2018) When Image Denoising Meets High-Level Vision Tasks: A Deep Learning Approach. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, Stockholm, 13-19 July 2018, 842-848. <https://doi.org/10.24963/ijcai.2018/117>
- [19] McCartney, E.J. (1976) Optics of the Atmosphere: Scattering by Molecules and Particles.
- [20] Narasimhan, S.G. and Nayar, S.K. (2000) Chromatic Framework for Vision in Bad Weather. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, Hilton Head, 15 June 2000, 598-605. <https://doi.org/10.1109/cvpr.2000.855874>
- [21] Silberman, N., Hoiem, D., Kohli, P. and Fergus, R. (2012) Indoor Segmentation and Support Inference from RGBD Images. *Computer Vision—ECCV 2012*, Florence, 7-13 October 2012, 746-760. https://doi.org/10.1007/978-3-642-33715-4_54
- [22] Lin, T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., *et al.* (2014) Microsoft COCO: Common Objects in Context. *Computer Vision—ECCV 2014*, Zurich, 6-12 September 2014, 740-755. https://doi.org/10.1007/978-3-319-10602-1_48