



张延军,陈博. 基于 HigherHRNet 的煤矿井下人体姿态估计快速网络研究[J]. 矿业安全与环保, 2024, 51(4): 35-40.

ZHANG Yanjun, CHEN Bo. Research on fast network of human pose estimation in underground coal mine based on HigherHRNet[J]. Mining Safety & Environmental Protection, 2024, 51(4): 35-40.

扫码阅读下载

DOI: 10.19835/j.issn.1008-4495.20240325

基于 HigherHRNet 的煤矿井下人体姿态估计 快速网络研究

张延军, 陈 博

(太原科技大学 机械工程学院, 山西 太原 030024)

摘要:煤矿井下人体姿态的快速估计是井下作业智慧安全检测的重要前提。为解决煤矿井下多尘多雾、照明不足及颜色相融等问题,提高人体姿态估计关键点分配准确度及网络运行速度,研究新的 Optimising HigherHRNet(OH-HRNet)快速网络模型:对 HigherHRNet 模型的轻量化设计、关键点分配进行深入研究,提出了基于注意力机制的记忆卷积模块及强化骨骼约束的关键点分配算法,并改进了算法的损失函数。在煤矿井下场景数据集和 COCO 公开数据集上的实验结果表明:OH-HRNet 在 GPU 的速度上是 LitePose 的 1.06 倍,平均精度均值 mAP 提高了 7.4%,平均召回率均值 mAR 提高了 14.0%,可以实现更有效的智慧安全检测。

关键词:智慧煤矿;人体姿态估计;快速网络;关键点检测;关键点分配

中图分类号:TD76

文献标志码:A

文章编号:1008-4495(2024)04-0035-06

Research on fast network of human pose estimation in underground coal mine based on HigherHRNet

ZHANG Yanjun, CHEN Bo

(School of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China)

Abstract: Fast estimation of human pose in underground coal mine is an important prerequisite for intelligent safety detection of underground operation. Aiming at the problems of dusty and foggy, insufficient illumination and color blending in underground coal mine, this study conducts an in-depth study on the lightweight design and key point assignment of the HigherHRNet model and proposes a new Optimising HigherHRNet(OH-HRNet) fast network model in order to improve the accuracy of the key point assignment for human pose estimation as well as the network operation speed. The OH-HRNet model proposes a memory convolution module based on attention mechanism and a key point assignment algorithm with reinforced skeletal constraints, and the loss function of the algorithm is improved. Experiments on the coal mine underground scenario dataset and COCO public dataset show that, OH-HRNet is 1.06 times faster than LitePose in terms of GPU speed, with a 7.4% increase in mAP and a 14.0% increase in mAR, which can achieve more effective intelligent safety detection.

Keywords: intelligent coal mine; human pose estimation; fast network; key point detection; key point assignment

在智慧矿山的建设过程中,安全问题始终是核

心关注点,相关领域的研究正在持续发展^[1-2]。据统计,2016—2022 年我国煤矿人的不安全行为导致的事故占比高达 50.5%^[3]。煤矿井下的生产环境复杂,多尘多雾、照明不足,同时人员误闯、违规操作等特殊场景也对人体姿态估计提出了新的挑战。另外,关键点检测不准确及延迟较长等问题,均会严重影响作业人员的实时反应和补救时间,进而易造成

收稿日期:2024-04-10;2024-05-23 修订

基金项目:山西省重点研发项目(202102010101010)

作者简介:张延军(1982—),男,河北丰宁人,教授,博士生导师,主要研究方向为智能控制。E-mail:hebei0735@126.com。

人员伤亡及严重的经济损失。因此,对井下人体姿态估计网络模型进行研究至关重要。

目前,国内外学者针对人体姿态估计网络开展了大量的研究工作。YU 等^[4]提出了 Lite-HRNet,将 ZHANG 等^[5]提出的 ShuffleNet 中的 Shuffle 块应用于 SUN 等^[6]提出的 HRNet,并提出一个高效加权单元;邓辉等^[7]提出了 LDANet,将 HAN 等^[8]提出的 GhostNet 的卷积结构应用于 HRNet 中,同时设计了密集单元和密集连接网络结构;BAJPAI 等^[9]提出了 MoveNet,借鉴 SANDLER 等^[10]提出的 Mobilenetv2 的卷积结构配合特征金字塔网络框架,实现快速人体姿态估计;OSOKIN^[11]提出的 Lightweight OpenPose,主要以 KIM 等^[12]提出的 OpenPose 作为框架,以 MobileNet 系列网络^[13-14]替换 VGG 特征提取器;林远强等^[15]提出的 Lite-CCA-HRNet 以 HRNet 为基础框架,利用 MetaFormer 结构和级联通道注意力来实现轻量化设计;马赛等^[16]提出了 LE-Higher-HRNet,利用深度可分离卷积和协调注意力机制实现人体姿态估计。然而,上述研究大多是针对井上人员开展的,而有关煤矿井下研究的报道甚少。

同时经研究发现,现有网络可以很好地完成人体姿态估计任务,但推理速度还有待提高。最初为追求更快速的神经网络,大多数网络利用深度可分离卷积^[17]和分组卷积^[18-19]减少神经网络的浮点运算量(FLOPs)。然而,网络运行速度还与内存的读写、网络设计的碎片化程度等因素相关^[20]。以 Lite-HRNet 为例,对比 HRNet_w32,虽然 FLOPs 降低至原来的 1/10,但推理时间是原来的 1.6 倍。

基于上述问题与研究,笔者设计了 Optimising HigherHRNet(简称“OH-HRNet”)模型,设计一种基于注意力机制的记忆卷积模块,探索基于部件关联场的强化骨骼关键点分配算法,强化骨骼对人体姿态的约束。构建煤矿井下人体数据集,同时利用 COCO 数据集,对 OH-HRNet 模型在煤矿井下场景和开放场景中人体姿态估计进行性能评估。

1 OH-HRNet 模型

HigherHRNet^[21]是由 HRNet 输出的特征图及通过上采样的高分辨率输出组成的,可应对自下而上多人姿态估计中的尺度变化挑战,并更精确地定位关键点位置。这也更符合煤矿井下人员分布情况,故以 HigherHRNet 作为基础框架,开展煤矿井下人体姿态估计算法研究。HigherHRNet 网络的具体结构如图 1 所示。

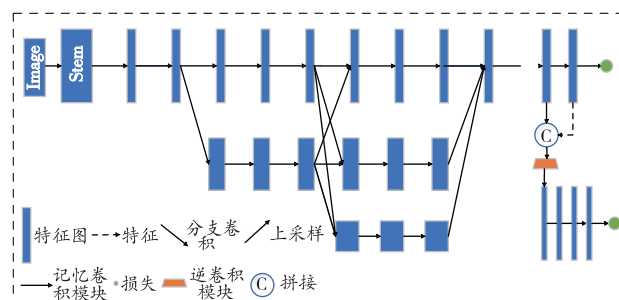


图 1 HigherHRNet 网络结构示意图
Fig. 1 Schematic diagram of the HigherHRNet structure

1.1 记忆卷积模块

数据分析时,把 2 次卷积后的 64 通道的特征图中每连续的 3 个通道的特征图合并为一张图片,并进行可视化处理。图 2 展示了不同通道的特征图(左上角为原图,其余为各通道特征图),可以看出各特征图之间有很大的相似度。

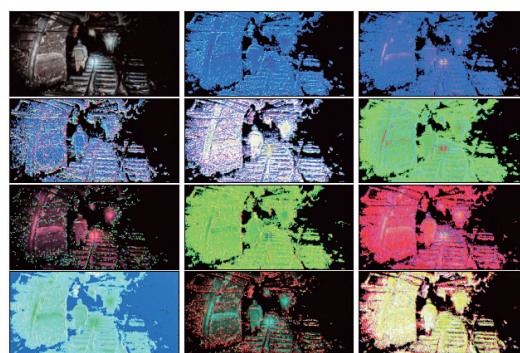


图 2 各卷积通道特征图
Fig. 2 Convolutional channel feature map

同时,经数据集分析发现,煤矿井下还存在环境多尘多雾、照明不足产生的算法噪声,以及员工长期工作导致工作服与煤矿背景的颜色相融等问题。此外,在井下独立光源条件中具有明显特征的工作服反光条,会引起头盔、煤矿设备和煤炭表面出现反光点,这些反光点会与反光条混淆,导致识别出现偏差,难以保证姿态估计效果,如图 3 所示。

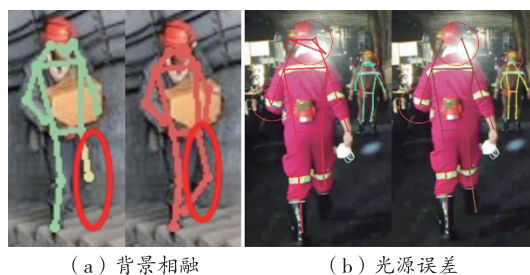


图 3 网络识别结果对比图(利用记忆增强模块改进前后效果)
Fig. 3 Comparison of network recognition results (improved results using memory enhancement module)

基于上述分析,设计了记忆卷积模块;以注意力机制为基础,分别就网络轻量化和关键点检测设计

了遗忘机制和记忆增强机制。

1.1.1 遗忘机制

采用 SENet 注意力机制^[22]对特征图的重要性进行排序,其工作原理如下:①利用 F_{sq} 函数和 F_{ex} 函数获取每个通道的权重向量,进行范数运算,得到对应通道特征图的记忆重要性,并作好标记;②根据需要将部分不重要通道的特征图分离出来,称之为遗忘特征图 F_o (Feature map-oblivion),对其不做处理,剩余的特征图即为记忆特征图 F_m (Feature map-memory)。遗忘机制具体结构如图 4 所示。

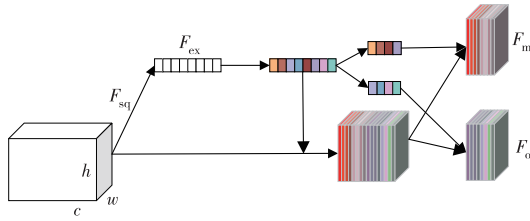


图 4 遗忘机制结构图

Fig. 4 Structure of the forgetting mechanism

根据具体的记忆率决定具体遗忘的特征图比例,减少网络计算量,达到加速目的。权重向量计算公式如下:

$$F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) = z \quad (1)$$

$$F_{ex}(z, W) = \delta(g(z, W)) = \delta(W_2 \delta(W_1, z)) \quad (2)$$

式中: u_c 为特征图输入; $H \times W$ 为特征图尺寸; W_1 为与 z 相匹配的一个尺寸(只有相匹配才能相乘), W_2 同理。

1.1.2 记忆增强机制

为更全面提取特征图信息,对 F_m 先进行逐通道卷积,并与原记忆特征图直接叠加,再通过多分辨率特征图的融合实现记忆加强,最后把遗忘特征图与加强后的特征图拼接在一起,通过点卷积对所有通道信息进行交流。记忆卷积模块结构如图 5 所示。

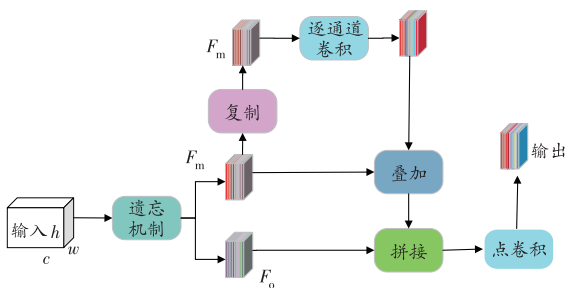


图 5 记忆卷积模块结构图

Fig. 5 Structure of memory convolution module

1.2 关键点分配算法

HigherHRNet 算法在煤矿井下的实际运行中,出现了关键点错分现象(不同颜色的连接线代表不同个体的肢体),如图 6(a) 所示。同时,为保证平均精

度均值 mAP,采用了尽可能将所有的关键点都标记出来的策略,使得平均召回率均值 mAR 较低,且某些本不存在的关键点却被误检,这也对关键点分配造成了一定的干扰(图 3(b))。



(a) 识别错误示意图 (b) 算法约束改进后识别效果图

图 6 关键点识别效果对比图

Fig. 6 Key point assignment recognition effect contrast

1.2.1 算法约束

针对上述问题,基于部件关联场,提出一种强化骨骼约束的关键点分配算法。利用人体骨骼的约束,提出 3 条新约束:

1) 骨骼长度约束。由于数据集使用的是 2D 平面图,角度问题可能导致关节点之间距离极短,故而提出了新的骨骼约束参数 ε : 任何 2 个关节点之间的距离不得超过其余关节点之间最大距离的 ε 倍,其数学公式如下:

$$l_{ij_2} \leq \varepsilon \max l_i \quad (3)$$

式中: l_{ij_2} 为同一肢体的 2 个关节点之间的距离; l_i 为第 i 段关键点之间的距离。

2) 骨骼曲度约束。在部件关联场提出了不得超出肢体宽度,但只留下一个手动调控的参数。在此基础上提出了骨骼约束: 对四肢关节点之间的位置角度提出限制,使 2 个关节点轮廓走向的角度不超过 25° ,其数学表达式如下:

$$\left| \frac{w_{j_2,k} - w_{j_1,k}}{x_{j_2,k} - x_{j_1,k}} \right| \leq \cos 25^\circ \quad (4)$$

式中: w 为某一肢体的名称; k 为人的编号; x 为 2 个相邻关节点之间的肢体点。

通过肢体点与关节点之间的角度约束,进一步强化各关节点之间的联系。

3) 关键点检测约束。在周围关键点没有与之匹配的情况下(即孤立关键点),该关键点的检测分数也会得到抑制,具体如下:

$$Q(j) = \begin{cases} q, & \text{if } z_{j,j_1}^{mn} = 1 \\ \varphi q, & \text{otherwise} \end{cases} \quad (5)$$

式中: $Q(j)$ 为关节点的置信度; $z_{j,j_1}^{mn} \in \{0,1\}$,为 2 个关节点之间是否连接(1 表示连接,0 表示未连接); m,n 为任意 2 个关节点; φ 为抑制参数,根据不同的应用环境可自行调节。

1.2.2 定义损失函数

首先假定点 p 为人体某一肢体 h 上的点,由于骨骼约束,导致该点到端点的向量与该肢体两端关键点向量方向一致,则有:

$$p(u) = \frac{x_{j_2,h} - x_{j_1,h}}{\|x_{j_2,h} - x_{j_1,h}\|_2} \quad (6)$$

$$E = \int_{u=0}^1 p(u) du \quad (7)$$

$$\text{Loss} = \frac{1}{|E|} \quad (8)$$

式中: $p(u)$ 为随机点 u 到肢体端点关键点的向量; E 为整条肢体上所有点到肢体端点的关键点之间的向量之和, E 越大,则对应肢体上的点到肢体端点关键点的方向越一致。

肢体的可信度通过求累加向量的模得出。

2 实验验证与分析

进行消融实验时,首先评估 OH-HRNet 结构在煤矿井下人体姿态估计任务中的表现,对网络进行泛化实验,针对 COCO 数据集开放场景进行人体姿态估计对比实验。

2.1 数据集及评估标准

数据集主要来源于煤矿井下和相关网络视频截

图,共包含 1 000 张图片,近 3 000 个标注人体。标注格式、评价标准和 COCO 数据集一样。同时,公开数据集采用 COCO 2017 主流的大规模人体关键点数据集,其包含有 80 个目标类目,多达 33 万张图片,20 万个标注的人体。

COCO 数据集关键点检测的评价指标采用平均精度 AP (average precision) 和平均召回率 AR (average recall) 进行评测。通过定义对象关键点相似度 OKS (object keypoint similarity) 来度量预测姿态和真实姿态的匹配相似性,并计算 AP 在 OKS 取值范围(0.50 : 0.05 : 0.95)的平均精度均值 mAP 作为主要的衡量指标。

2.2 实验设置

数据增强包括随机旋转 $[-30^\circ, 30^\circ]$,随机尺度变换 $[0.75, 1.50]$,随机平移 $[-40, 40]$,以及随机翻转。

在 COCO 数据集上进行了 180 个 epoch 的训练,设置 batch size = 32,学习率初始值为 0.005,学习率调度器 (Learning Rate Scheduler) 选用 ReduceLROnPlateau,其中 LR_FACTOR: 0.8, LR_PATIENCE: 5。

2.3 消融实验

为验证所设计网络的必要性和有效性,以 HigherHRNet 为基础开展消融实验,分别对比记忆卷积模块、强化骨骼约束不同参数对网络的影响。实验采用 COCO 2017 (2017 版 COCO) 数据集进行验证,分别取记忆率为 1/2、1/4 进行训练,对比其推理速度与 mAP,结果见表 1。

表 1 消融实验对比
Table 1 Comparison table of ablation experiments

网络结构	记忆卷积模块		记忆率 r	强化骨 骼约束	尺度约 束 ε	关键点 约束 φ	卷积核 尺寸	吞吐量/ (Samples · s ⁻¹)	延迟/ s	mAP/ %	mAR/ %
	遗忘机制	记忆增强									
HigherHRNet	×	×	×	×	×	×	3	1.2	0.833	67.1	69.4
HigherHRNet_o	√	×	1/2	×	×	×	3	24.4	0.041	35.8	36.5
			1/4	×	×	×	3	25.6	0.039	36.0	36.7
HigherHRNet_m	×	√	1/4	×	×	×	3	22.3	0.044	54.6	53.8
OH-HRNet	√	√	1/4	√	1.5	0.8	3	18.6	0.054	50.7	58.7
			1/4	√	1.5	0.9	3	19.7	0.051	53.3	59.6
			1/4	√	2.0	0.8	3	18.4	0.054	50.6	57.3
	√	√	1/4	√	2.0	0.9	3	19.9	0.050	53.2	59.8
			1/4	√	1.5	0.9	5	18.2	0.055	57.0	63.6
			1/4	√	1.5	0.9	7	17.6	0.057	61.7	67.6

注:HigherHRNet_o 为只包括遗忘机制的算法;HigherHRNet_m 为包含了记忆增强的算法。

由表 1 可知,当记忆率 $r=1/2$ 时,遗忘机制使网络速度(吞吐量)提高为原来的 20.3 倍;当记忆率 $r=1/4$ 时,遗忘机制使网络速度提高为原来的 21.3 倍,且准确率要比 $r=1/2$ 时高 0.2%,故在后续的实验中,设置记忆率 $r=1/4$ 。此外,记忆增强模块使网络的准确率提高了 18.6%,而延迟只增加了 0.05 s。

在强化骨骼算法中,对尺度约束 ε 和关键点约束 φ 2 个参数进行了简单的对比实验(ε 分别取 1.5、2.0, φ 分别取 0.8、0.9)。当 $\varepsilon=1.5$ 、 $\varphi=0.9$ 时,随着关键点约束参数的设置,图像中关键点检测的准确率降低了 1.3%,但是召回率提高了 5.8%,这

也意味着网络对于关键点的检测更准确可靠。故在后续的实验中,以 $\varepsilon=1.5$ 、 $\varphi=0.9$ 作为约束参数。

分别对比了卷积核尺寸为 3、5、7 时网络的 mAP,实验结果表明大卷积核可更有效地实现对空间信息和语义特征的提取。卷积核在 mAP 上,尺寸为 7 的要比尺寸为 3 的提高 8.4%。故在后续的实验中,均将卷积核大小为 7 作为实验参数。

2.4 煤矿井下实验

将 OH-HRNet 与现有快速人体姿态估计算法进行对比实验。数据集为煤矿井下的数据集,输入 Input size 为 512 像素,批量大小为 32 的各网络实验结果见表 2。

表 2 煤矿井下场景各网络准确率、推理延迟
Table 2 Accuracy, inference delay of each network in coal mine underground scenario

网络类别	特征图尺寸/ (像素×像素)	计算量/ GFLOPs	GPU		CPU		mAP/%	mAR/%
			吞吐量/ (Samples · s ⁻¹)	延迟/s	吞吐量/ (Samples · s ⁻¹)	延迟/s		
Lightweight OpenPose ^[11]	368×368	4.5	10.68	0.093	0.356 28	5.657 2	35.8	37.9
EfficientHRNet ^[23]	448×448	4.0	9.43	0.106	0.375 18	5.529 5	42.7	46.0
LitePose ^[24]	448×448	2.5	16.52	0.061	0.434 57	4.641 0	49.6	50.9
OH-HRNet	512×512	12.1	17.60	0.057	0.452 97	4.146 8	57.0	64.9

由表 2 可以看出,改进设计的 OH-HRNet 在煤矿井下场景中的性能是最优的。该网络成功实现了性能—计算权衡的最优解,在保证速度的同时大大提高了煤矿井下人体姿态估计的识别精度。在 GPU 上相比于之前的 LitePose (STOA: LitePose), mAP 提高了 7.4%, mAR 提高了 14.0%。同时速度上是 LitePose 的 1.06 倍。与 Lightweight OpenPose 相比,更是实现了速度近 1.65 倍的突破, mAP 提高了 21.2%。具体识别效果如图 7 所示。



图 7 煤矿井下人体姿态估计效果图
Fig. 7 Effect of humanpose estimation in coal mine underground

3 结束语

1) 针对煤矿井下复杂环境下的人体姿态估计网络速度与精度的平衡问题及关键点分配问题,提出一种 OH-HRNet 模型。消融实验表明:在 COCO 数据集中,遗忘机制使网络速度提高为原来的 21.3 倍,记忆增强机制可以使 mAP 提高到 54.6%,使用 7×7 卷积核可以使 mAP 提高到 61.7%,强化骨骼约束分配算法可以使召回率提高到 67.6%。

2) 在煤矿井下人体姿态方面,OH-HRNet 模型在 CPU 和 GPU 运行速度上,分别是网络 LitePose 速度的 1.04 倍和 1.06 倍, mAP 提高了 7.4%, mAR 提高了 14.0%。

参考文献 (References):

[1] 王海钢. 煤矿 5G 智能化应用实践关键技术研究[J]. 矿业安全与环保, 2023, 50(5): 94-98.
WANG Haigang. Research on key technologies of 5G intelligent application practice in coal mine[J]. Mining Safety & Environmental Protection, 2023, 50(5): 94-98.
[2] 徐丽丽, 栗继祖, 徐新华. 智能化矿井监控调度作业的认知作业分析及认知疲劳判定[J/OL]. 矿业安全与环保: 1-7 [2024-05-23]. <http://kns.cnki.net/kcms/detail/50.1062.TD.20240229.0911.002.html>.

- XU L L, LI J Z, XU X H. Cognitive task analysis and cognitive fatigue determination for intelligent[J/OL]. Mining Safety & Environmental Protection: 1-7[2024-05-23]. <http://kns.cnki.net/kcms/detail/50.1062.TD.20240229.0911.002.html>.
- [3] 程磊, 孙洁. 2016—2022年我国煤矿事故统计与规律分析[J]. 煤炭工程, 2023, 55(11): 125-129.
- CHENG Lei, SUN Jie. Statistics and law analysis of coal mine accidents in China from 2016 to 2022[J]. Coal Engineering, 2023, 55(11): 125-129.
- [4] YU C Q, XIAO B, GAO C X, et al. Lite-HRNet: A lightweight high-resolution network[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA. IEEE, 2021: 10435-10445.
- [5] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 6848-6856.
- [6] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA. IEEE, 2019: 5686-5696.
- [7] 邓辉, 徐杨. 融入注意力和密集连接的轻量化人体姿态估计[J]. 计算机工程与应用, 2022, 58(16): 265-273.
- DENG Hui, XU Yang. Lightweight human pose estimation based on attention and dense connection[J]. Computer Engineering and Applications, 2022, 58(16): 265-273.
- [8] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020: 1577-1586.
- [9] BAJPAI R, JOSHI D. MoveNet: A deep neural network for joint profile prediction across variable walking speeds and slopes[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 2508511.
- [10] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 4510-4520.
- [11] OSOKIN D. Real-time 2D multi-person pose estimation on CPU: Lightweight OpenPose[EB/OL]. 2018: ArXiv:1811.12004.
- [12] KIM W, SUNG J, SAAKES D, et al. Ergonomic postural assessment using a new open-source human pose estimation technology (OpenPose)[J]. International Journal of Industrial Ergonomics, 2021, 84: 103164.
- [13] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. 2017: ArXiv:1704.04861.
- [14] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South). IEEE, 2019: 1314-1324.
- [15] 林远强, 郜辉, 王鹏, 等. 引入级联通道注意力的轻量化人体姿态估计[J]. 计算机工程与应用, 2024, 60(13): 219-227.
- LIN Y Q, GAO H, WANG P, et al. Lightweight human pose estimation with cascaded channel attention[J]. Computer Engineering and Applications, 2024, 60(13): 219-227.
- [16] 马赛, 葛海波, 何文昊, 等. 轻量高效的自底向上人体姿态估计算法研究[J/OL]. 计算机工程与应用, 1-22[2024-04-29]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20230814.1802.022.html>.
- MA S, GE H B, HE W H, et al. Research on lightweight and efficient bottom-up human pose estimation algorithm[J/OL]. Computer Engineering and Applications. 1-22[2024-04-29]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20230814.1802.022.html>.
- [17] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 1800-1807.
- [18] SIFRE L, MALLAT S. Rigid-motion scattering for texture classification[EB/OL]. 2014: ArXiv:1403.1687. <http://arxiv.org/abs/1403.1687>.
- [19] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [20] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNetV2: practical guidelines for efficient CNN architecture design[C]//European Conference on Computer Vision. Cham: Springer, 2018: 122-138.
- [21] CHENG B W, XIAO B, WANG J D, et al. HigherHRNet: Scale-aware representation learning for bottom-up human pose estimation[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020: 5385-5394.
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [23] NEFF C, SHETH A, FURGURSON S, et al. EfficientHRNet[J]. Journal of Real-Time Image Processing, 2021, 18(4): 1037-1049.
- [24] WANG Y H, LI M Y, CAI H, et al. Lite pose: Efficient architecture design for 2D human pose estimation[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA. IEEE, 2022: 13116-13126.

(责任编辑: 林桂玲)