

面向智慧养老的跌倒判别边缘计算终端设计与实现*

刘芳宇, 范云, 祁瑞安, 黄威

(内蒙古大学电子信息工程学院, 呼和浩特 010070)

摘要:数字经济与人工智能深度融合推动智慧城市建设进程持续加快,智慧养老作为智慧城市民生服务的重要支撑板块,其对场景内信息传输提出了更高要求。针对当前跌倒检测存在的信息传输与响应滞后以及隐私安全隐患突出两大核心问题,设计并实现一套跌倒姿态判别硬件系统。采用边缘计算架构,以改进YOLOv11-Pose为核心,通过MobileNetV4优化主干网络降低计算复杂度,同时制定融合骨骼点特征的多维度评分算法,实现跌倒行为识别与跌倒方向判别。算法部署于Jetson Nano硬件终端,在边缘端即时响应,大幅减少数据传输量,规避隐私泄露风险。实验结果表明,该方法能够精准区分不同姿态并正确判别跌倒方向。边缘硬件终端实测结果显示,系统判别延迟低、运行稳定,可为智慧养老场景提供高效、可靠的技术支撑。

关键词:边缘计算; 人体姿态估计; 跌倒方向判别; YOLOv11-Pose

中图分类号:TN915 **文献标志码:**A

在数字经济与人工智能深度融合的背景下,智慧城市建设已成为城市发展的核心方向。智慧养老作为智慧城市民生服务体系的重要组成部分,其安全保障能力直接关系到智慧城市建设的民生价值落地^[1]。跌倒是导致老年人重伤、失能的主要诱因,因此,适配智慧养老需求的跌倒检测技术已成为支撑智慧城市养老服务体系的核心环节^[2]。随着家用监控普及和计算机视觉技术的进步,从视频图像中提取人体特征实现跌倒检测逐渐成为主流,依托智能化通信技术实现跌倒判别信息的即时传输,同时保障老年人的个人隐私,是跌倒检测技术发挥实效、切实守护老年人生命安全的关键前提^[3]。

目前的跌倒检测方法主要分为非视觉检测方法和视觉检测方法。非视觉检测方法主要包括可穿戴传感器检测、环境式传感器检测与毫米波雷达检测3种技术路径。可穿戴传感器检测算法通过将传感器设备采集的信号转换为代表运动状态的信息,从而判断是否发生跌倒^[4]。Alarifi等^[5]将磁力计、陀螺仪和加速度计组成的设备放置在受试者身体的6个不同位置进行跌倒检测。虽然该方法具有较高的识别精度,但使用复杂且舒适性较差。环境式传感器主要布设于易跌倒的特定区域,提供发生跌倒的位置信息。Yazar等^[6]提出集成1个振动传感器和2个被动红外传感器,实现跌倒检测、脚步声检测、人体运动检测及异常静止检测。这种方法场景迁移性差且易受环境干扰导致误

* 收稿日期:2025-11-03; 修回日期:2026-04-03

基金项目:国家自然科学基金项目(62561043);内蒙古自治区重点研发和成果转化计划项目(2023YFDZ0056);
鄂尔多斯市重点研发计划项目(YF20250282)

作者简介:刘芳宇(2001—),女,内蒙古包头人,2023级硕士研究生。E-mail:liufangyu0606@163.com

通信作者:黄威(1978—),女,内蒙古呼伦贝尔人,教授,硕士。主要从事无线通信及其应用、FPGA硬件加速研究。E-mail:huangwei@imu.edu.cn

报。近年来,毫米波雷达因其非接触式检测特性被广泛应用于跌倒检测领域。Xu等^[7]提出毫米波连续运动检测系统,通过发射毫米波信号捕获人体运动轨迹与姿态变化,将毫米波雷达的微多普勒信号转化为可视化特征,结合目标检测技术实现运动检测与跌倒行为识别。但毫米波雷达对复杂居家环境中人体细微动作的识别精度较低,易受家具和墙体等障碍物的遮挡,从而影响其检测稳定性。

视觉检测算法通过图像采集设备获取画面,结合姿态分析或行为识别模型判断是否发生跌倒^[8]。Min等^[9-10]将人体边界框宽高比变化作为特征进行跌倒检测。Xu等^[11]借助OPENPOSE形成的人体骨架图进行姿态识别。在此基础上,Chen等^[12]引入髋关节中心下降速度、人体中心线与地面的夹角、人体外接矩形宽高比3个关键参数识别跌倒事件。此外,Lou等^[13]提出了基于摄像头的YOLOv8的小目标检测算法,通过深度可分离卷积的下采样操作提取特征信息,同时改进原有模型的C2f模块进行不同尺寸的特征融合,实现对小目标检测算法性能的提升。基于视觉的跌倒检测算法融合了多种参数以提升精度,但复杂的网络结构会受到边缘设备算力的限制,无法满足移动和嵌入式设备的部署需求。

综上所述,传统跌倒检测技术难以适配智慧养老体系对“即时响应+隐私保护”的双重需求。可穿戴传感器存在佩戴舒适性差、老年人抵触率高的问题,环境式传感器场景迁移性弱,毫米波雷达虽能规避隐私风险,但检测精度受遮挡影响大。视觉检测方法虽具备场景适应性优势,但需传输全量图像数据至远端终端处理,在多设备并发场景下受限于传输带宽,数据传输与云端处理总延迟远高于跌倒后“1分钟黄金救助时间”的即时性要求,且视觉检测方案包含老年人居家环境细节、身体特征的完整图像,这些数据在传输至远端终端的过程以及在云端存储时,存在被泄露、滥用的风险,既侵犯老年人隐私,也不符合智慧养老场景下的安全管理规范。除此之外,目前的跌倒检测算法仅能判断是否发生跌倒事件,缺乏对跌倒方向信息的判别,难以支撑智慧养老场景下的精准救治调度,进一步制约了跌倒检测技术在智慧养老场景中的实际应用效果。

针对上述瓶颈,以边缘计算为核心技术架构,将图像采集、特征提取、姿态判别全流程部署于边缘硬件终端,仅向云端网络传输跌倒状态的结果数据,而非全量图像。既减少通信带宽占用,将响应延迟缩短,满足即时预警需求,又从源头保护了老年人的隐私安全。由于边缘硬件终端算力有限,通过轻量化算法改进适配边缘算力,在保障判别精度的同时提升推理速度,确保终端能稳定处理图像数据。首先,使用MobileNetV4^[14]改进YOLOv11-Pose原有主干网络,并引入混合注意力机制(Shuffle attention, SA)^[15]作为特征提取网络,提升模型对复杂环境下骨骼关键点的识别能力。其次,利用人体外接矩形框宽高比和骨骼点融合特征作为判别条件,设计评分规则进行得分计算,动态调整姿态估计得分,并对跌倒姿态进一步判别跌倒方向。最后,将改进后的模型在自建数据集上进行测试,并部署于硬件终端模拟居家养老场景中的实时检测。实验结果证明,改进算法能在快速识别人体姿态的同时准确判别跌倒方向,为智慧养老领域提供了有效的技术支持。

1 人体骨骼关键点检测算法

YOLOv11-Pose作为当前主流的人体姿态估计模型,其网络架构由主干网络、颈部网络与头部网络3部分组成。主干网络采用C3k2模块作为核心特征提取单元,通过多组C3k2模块堆叠实现从图像中提取浅层到深层的特征信息,其中CBS模块为卷积、归一化与激活函数SiLU的结合。快速空间金字塔池化(SPPF)模块包含2个CBS模块和3个最大池化层,通过采用不同尺寸的池化核,在捕捉输入图像多尺度信息的同时减少通道数,增强模型对于不同尺度特征的感知能力。颈部网络通过特征金字塔结构与路径聚合网络,实现主干网络输出特征的进一步整合与传递,提升特征的语义表达能力。头部网络通过不同预测分支输出人体外接矩形框坐标与17个骨骼关键点位置信息。

由于YOLOv11-Pose网络模型参数量大、计算复杂度高,受到硬件资源和算力的限制,难以直接

部署于硬件终端。图1为改进后模型的整体网络结构,改进网络使用MobileNetV4核心模块通用倒置瓶颈(Universal inverted bottleneck, UIB)替代主干网络中原有C3k2模块以减少参数量,在此基础上引入SA保证定位信息的准确性。

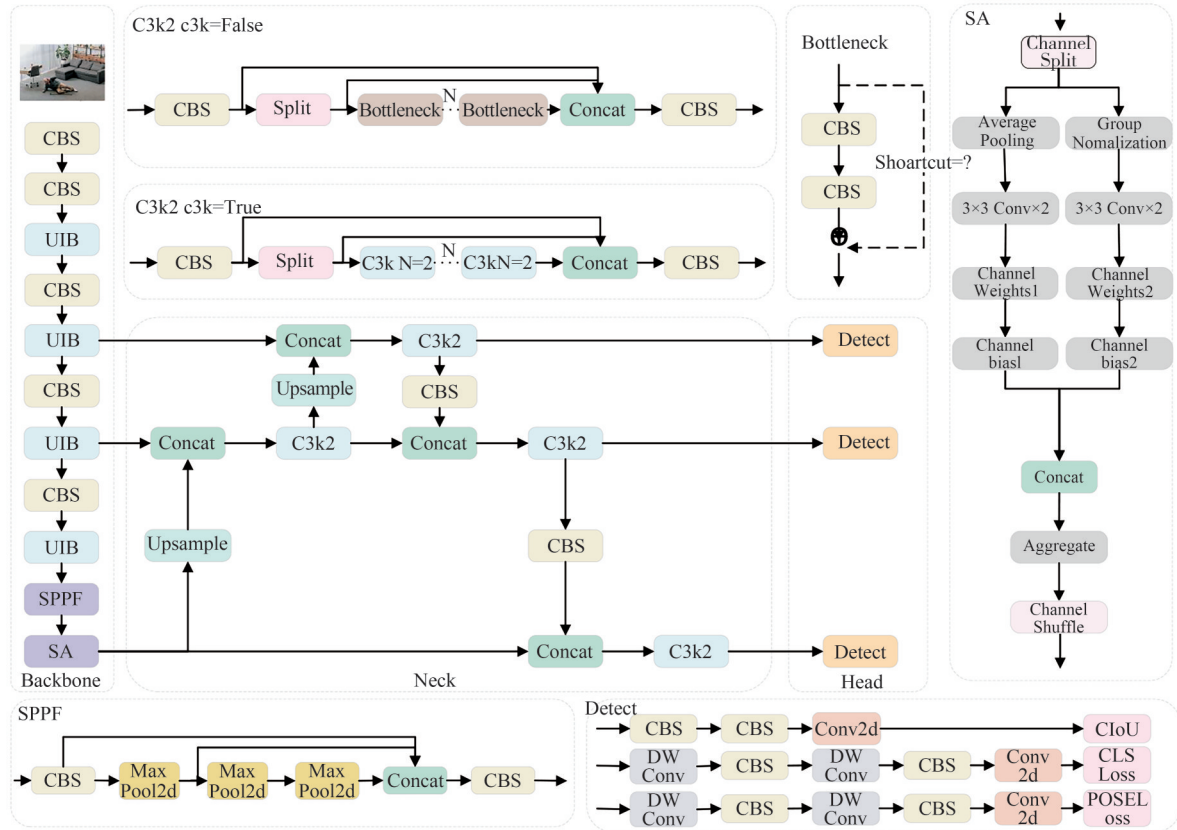


图1 改进型YOLOv11-Pose网络结构

Fig. 1 Improved YOLOv11-Pose network structure

MobileNetV4凭借其高效架构设计,有效减少了计算量和参数量,在保证检测精度的同时,显著提升了模型推理速度,以适配不同硬件平台。通过引入通用UIB搜索块和移动端多查询注意力(Mobile MQA)模块,优化神经网络架构的搜索方法,大幅提升模型运行效率。将普通 3×3 卷积拆分为深度可分离卷积和逐点卷积是模块改进的核心要素。图2为UIB模块结构。

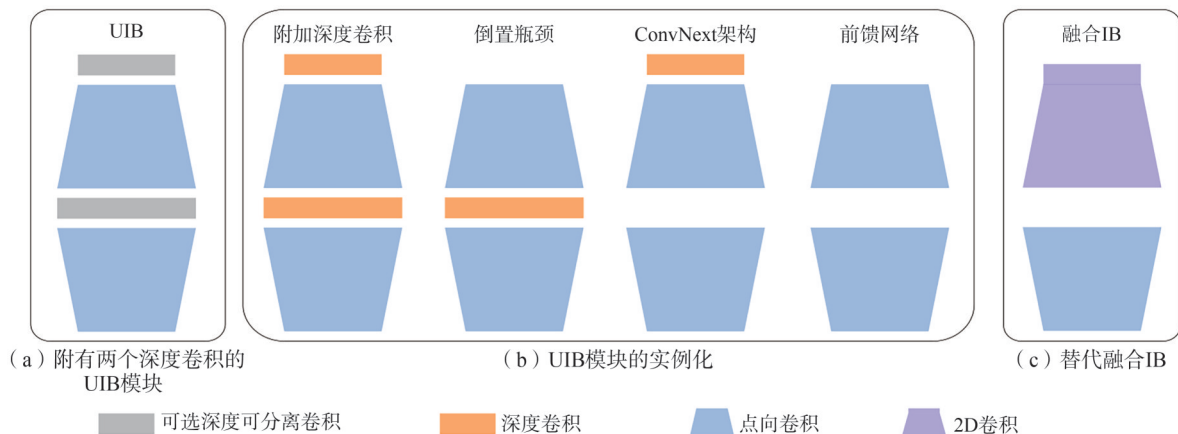


图2 通用倒置瓶颈结构

Fig. 2 Universal inverted bottleneck structure

UIB引入两个可选深度可分离卷积层,一个在扩展层之前,另一个在扩展层和投影层之间,有效统一了MobileNet网络中的原始模块。图2(b)为UIB模块中可选深度卷积4种可能的实例化,在每个网络阶段,UIB提供灵活的实例化策略,实现即时的空间和通道混合平衡,在平衡计算开销的同时按需扩大感受野,实现计算资源的高效利用。

为强化人体骨骼关键点特征,减少智慧养老复杂居家环境中家具和光线等干扰因素的影响,准确捕捉人体轮廓的微弱变化,更好地区分老年人的日常生活行为和意外跌倒行为,在改进网络中引入混合注意力机制,其网络结构如图3所示。

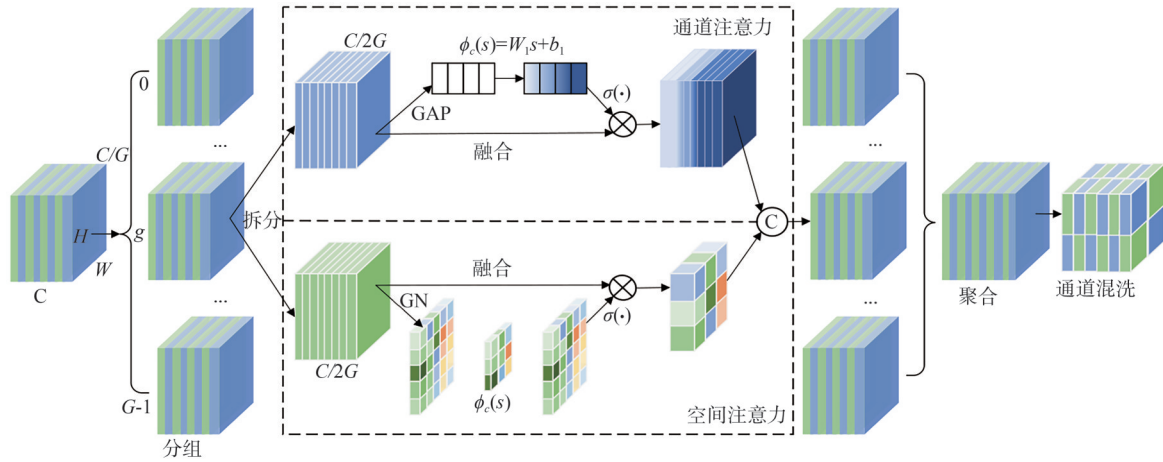


图3 混合注意力机制

Fig. 3 Shuffle attention

将主干网络得到的尺寸为 $C \times H \times W$ 的特征图 $K (K \in \mathbb{R}^{C \times H \times W})$ 分成 G 个子特征组 $K (K = [K_1, K_2, \dots, K_G], K \in \mathbb{R}^{(C/G) \times H \times W})$, 然后进行并行处理, 此时每一个子特征组的通道数为 C/G , 将第 g 组的子通道进行分离, 重新得到两个新的分支 K_{g1} 和 K_{g2} , ($K_{g1}, K_{g2} \in \mathbb{R}^{(C/2G) \times H \times W}$), 每个分支的通道数为 $C/2G$ 。其中, g_1 分支将尺寸为 $H \times W$ 的特征图中的像素 $R_{g1}(i, j)$ 进行全局平均池化 (Global average pooling), 并生成通道统计 s , 表示为

$$s = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W R_{g1}(i, j)。$$

对 s 的通道向量进行参数缩放 W_1 和移动偏置 b_1 , 实现 $\sigma(\cdot) = \text{sigmoid}(\cdot)$ 平滑, 同时与 $R_{g1}(i, j)$ 对应位置元素进行乘积, 得到新的特征图像素 R'_{g1} , 公式为 $R'_{g1} = \sigma(W_1(s) + b_1) R_{g1}$ 。

在第 g 组的子特征分别实现通道注意力机制和空间注意力机制处理后, 将两个分支信息进行连接, 得到 R'_g , 公式为 $R'_g = [R'_{g1}, R'_{g2}]。$

在所有分组的子特征都完成注意力监督与特征提取后, SA 模块对各特征组进行通道混洗, 从而实现跨组的特征交互与信息传递。

SA 不仅强化了模型的特征表达能力, 提升了网络对关键特征的提取效率与定位精度, 还有效缓解了深层网络训练中易出现的梯度消失或梯度爆炸问题, 保障了模型训练的稳定性。本文采用 SA 模块替换 YOLOv11-Pose 关键点检测模型中跨阶段部分极化自注意力模块 (C2PSA) 的常规注意力机制, 在保持注意力效果的前提下, 借助特征分组与通道置换策略, 有效降低了模型的整体计算复杂度, 实现了精度与实时性的平衡。

2 融合骨骼关键点姿态判别

多维度评分规则能够全面覆盖智慧养老场景下的人体姿态, 为精准判别跌倒行为提供可靠的方

法支撑。

2.1 人体姿态估计

如图4所示,算法共检测出17个人体关键点,精准量化老年人各部位相对位置关系,为智慧养老场景下区分站立、坐姿、跌倒姿态提供细节支撑。算法分别对各关键点进行0—16编号。为了得到骨骼关键点的变化信息,以水平向右为*x*轴正方向、竖直向下为*y*轴正方向建立平面直角坐标系,关键点的编号用变量*p*=(*p*₀,*p*₁,⋯,*p*₁₆)表示,每一个关键点*p*定义其坐标为*p*_{*i*}=(*x*_{*i*},*y*_{*i*})。

利用得到的人体骨骼点位置信息计算骨骼点间中心坐标,包括肩膀、胯部、膝盖和脚踝等。以肩部中心点为例,左肩对应坐标(*x*₅,*y*₅),右肩对应坐标(*x*₆,*y*₆),则肩部中心点为*s*_{*c*}(*X*_{*s*_{*c*}},*Y*_{*s*_{*c*}}),其中,*X*_{*s*_{*c*}}=(*x*₅+*x*₆)/2,*Y*_{*s*_{*c*}}=(*y*₅+*y*₆)/2。

定义人体中心线为肩部中心坐标*s*_{*c*}和胯部中心坐标*h*_{*c*}的连线,确定(0,0)(10,0)连线为*x*轴上从原点开始的水平线,身体倾斜角度为*θ*,则

$$\cos \theta = \frac{x_{s_c} - x_{h_c}}{\sqrt{(x_{s_c} - x_{h_c})^2 + (y_{s_c} - y_{h_c})^2}}, \theta = \arccos \left(\frac{x_{s_c} - x_{h_c}}{\sqrt{(x_{s_c} - x_{h_c})^2 + (y_{s_c} - y_{h_c})^2}} \right)。$$

人体矩形框包含两个坐标值(*x*₀,*y*₀)(*x*₁,*y*₁),代表检测框的左上和右下坐标。人体宽高比定义为*a*_{*r*}=*w*/*h*,其中*w*=*x*₁−*x*₀,*h*=*y*₁−*y*₀。

*d*_{*v*}为肩部和膝盖在*y*轴方向的垂直距离差,公式为*d*_{*v*}=|*y*_{*s*_{*c*}}−*y*_{*k*_{*c*}}|,反映肩部和膝盖在竖直方向上的相对位置。其中,|·|表示取绝对值操作,保证得到的距离差为非负值。

设定垂直阈值进行判断,如果相对距离*d*_{*v*}值较小,表示肩部和膝盖几乎在同一水平面,为站立姿态;*d*_{*v*}值较大则表示人体存在倾斜,处于跌倒姿态。以相同计算方式计算肩部和膝盖在*x*轴方向的水平距离差*d*_{*h*},并设定水平阈值进行判断。

以肩部*s*_{*c*}、胯部*h*_{*c*}和膝盖中心点*k*_{*c*}形成夹角为例,判断腿部与身体上半身的位置关系。通过欧几里得距离公式计算边长,通过余弦定理计算*s*_{*c*}、*h*_{*c*}和*k*_{*c*}形成的角度为*θ*_{*s*_{*c*}*h*_{*c*}*k*_{*c*}}=arccos((*d*_{*s*_{*k*}}²−*d*_{*s*_{*h*}}²−*d*_{*h*_{*k*}}²)/−2*d*_{*s*_{*h*}}*d*_{*h*_{*k*}}),其中*d*=√(*x*₂−*x*₁)²+(*y*₂−*y*₁)²。

以相同计算方式计算胯部*h*_{*c*}、膝盖*k*_{*c*}和脚踝中心点*a*_{*c*}之间的夹角*θ*_{*h*_{*k*}*a*},进一步分析腿部姿势。

本文设计一套评分规则,将身体倾斜角度、人体外接矩形框宽高比、肩膀-膝盖垂直距离与水平距离、肩膀-胯部-膝盖夹角以及胯部-膝盖-脚踝夹角等作为判断条件,对站立、跌倒、坐姿、跌坐以及其他姿态进行检测,每种姿态初始分数均设为0。表1分别对身体倾斜角度*θ*、宽高比*a*_{*r*}、垂直相对距离*d*_{*v*}和水平相对距离*d*_{*h*}进行阈值设置。

根据阈值进行规则制定,当身体倾斜角度*θ*在±25°范围内时,人体中心线几乎平行

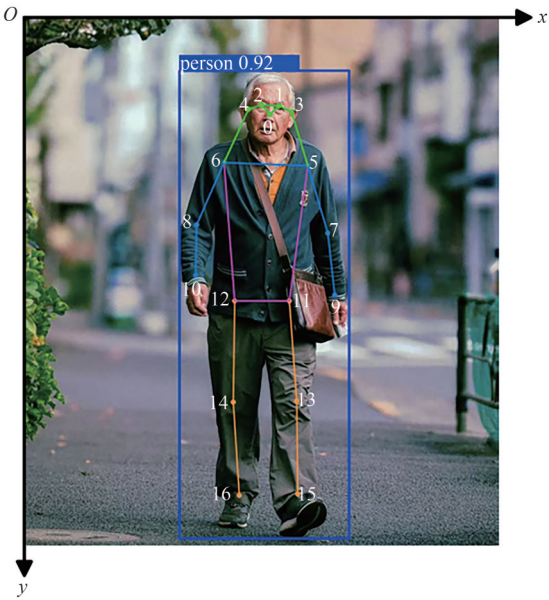


图4 人体骨骼关键点示意图
Fig. 4 Schematic diagram of human skeletal key points

表1 判别条件阈值	
Table 1 Discrimination condition threshold	
判别条件	阈值
身体倾斜角度 <i>θ</i> /(°)	25
宽高比 <i>a</i> _{<i>r</i>}	0.6
垂直相对距离 <i>d</i> _{<i>v</i>} /px	30
水平相对距离 <i>d</i> _{<i>h</i>} /px	30

于水平线,人体跌倒概率较高,跌倒姿态得分增加0.8。对于其他值,将根据与90°人体直立状态的差值增加对应跌倒的分数,角度偏离越大,跌倒的可能性越高。正常站立时人体轮廓窄而高,宽高比介于0.3~0.6之间,而跌倒时身体呈水平姿态,躯干与地面平行,轮廓宽高比显著增大。当宽高比小于设定阈值0.6且身体倾斜角度在65°到115°范围内时,表明人体处于站立状态,站立得分增加0.8;如果宽高比大于1/0.6,跌倒得分增加0.8。同样,垂直相对距离小于阈值30,表明膝盖和肩部在y轴上距离位置接近,跌倒得分增加0.6。当水平相对距离大于30,说明膝盖和肩部在x轴上距离过大,跌倒得分增加0.6,坐姿得分减小。跌坐姿态作为跌倒和坐姿的混淆姿态,其差别在于跌坐姿态身体上半身保持垂直,与膝盖夹角 $\theta_{s,h,k}$ 趋于90°,而腿部姿态 $\theta_{h,k,a}$ 接近平行,轮廓宽高比(1.6~1.8)介于站立与跌倒之间,故取值为1/0.6,即将1.67作为判定阈值,对应进行得分调整。姿态估计得分计算细节如表2所示,其中, S_{tand} 表示站立姿态, S_{it} 表示坐姿, F_{all} 表示跌倒姿态, F_{sit} 表示跌坐姿态。

表2 姿态估计得分计算表

Table 2 Calculation table for pose estimation score

姿态估计	判别条件	得分调整
身体倾斜角度 θ	$-25^\circ < \theta < 25^\circ$	$F_{\text{all}} + = 0.8, S_{\text{it}} + = 0.1$
	其他	$F_{\text{all}} + = 0.2 \times ((90 - \theta) / 90)$
宽高比 a_r	$a_r < 0.6, 65^\circ < \theta < 115^\circ$	$S_{\text{tand}} + = 0.6$
	$a_r > 1/0.6$	$F_{\text{all}} + = 0.8, F_{\text{sit}} + = 0.6$
垂直相对距离 d_v	$d_v < 30$	$F_{\text{all}} + = 0.6, S_{\text{it}} + = -0.15, F_{\text{sit}} + = 0.3$
水平相对距离 d_h	$d_h > 30$	$F_{\text{all}} + = 0.6, S_{\text{it}} + = -0.15, F_{\text{sit}} + = 0.3$
肩膀-胯部-膝盖之间夹角 $\theta_{s,h,k}$ 、身体倾斜角度 θ	$25^\circ < \theta_{s,h,k} < 145^\circ, 75^\circ < \theta < 125^\circ$	$F_{\text{sit}} + = 0.8, S_{\text{it}} + = 0.8, S_{\text{tand}} + = -0.035$
	$d_v > 30: S_{\text{it}} + = 0.15$	
	$120^\circ < \theta_{s,h,k}, 75^\circ < \theta < 125^\circ$	$S_{\text{tand}} + = 0.2$
	$120^\circ < \theta_{s,h,k}, -25^\circ < \theta < 25^\circ$	$F_{\text{all}} + = 0.2$
	其他	$F_{\text{all}} + = 0.05, S_{\text{tand}} + = 0.05$
胯部-膝盖-脚踝之间夹角 $\theta_{h,k,a}$ 、宽高比 a_r	$\theta_{h,k,a} > 125^\circ, a_r < 0.6$	$S_{\text{tand}} + = 1.6$
	$-25^\circ < \theta_{h,k,a} < 25^\circ, a_r > 1/0.6$	$F_{\text{all}} + = 1.6, F_{\text{sit}} + = 0.8$
	$75^\circ < \theta_{h,k,a} < 125^\circ$	$S_{\text{it}} + = 0.6$
胯部-膝盖-脚踝之间夹角 $\theta_{h,k,a}$ 、身体倾斜角度 θ	$25^\circ < \theta_{h,k,a} < 145^\circ, 75^\circ < \theta < 125^\circ$	$F_{\text{sit}} + = 0.3, S_{\text{it}} + = 0.8, S_{\text{tand}} + = -0.035$
	$d_v > 30: S_{\text{it}} + = 0.15$	
	其他	$F_{\text{all}} + = 0.05, S_{\text{tand}} + = 0.05$

通过组合关节角度和身体倾斜角度、宽高比进行判断,有效避免了相同身体倾斜角度对姿态估计的干扰,准确应对多样的姿态变化。完成人体姿态得分计算后,跌倒姿态得分最高时,将进入跌倒方向得分计算流程。跌倒方向依据受伤位置不同分为前向跌倒、侧向跌倒和后向跌倒。

2.2 跌倒方向判别

定义身体侧向差异为左、右肩膀在y轴上的竖直位置距离 d_s ,并检查与人体外接矩形框高度的相对偏移程度, R_s 为肩膀偏移量,公式为 $d_s = |y_{s,l} - y_{s,r}|$, $R_s = d_s/h$,其中, $y_{s,l}$ 为左肩对应y轴坐标, $y_{s,r}$ 为右肩对应y轴坐标。

定义肘部与肩部的相对偏移量 $R_{e,s}$ 为肘部与肩部在y轴上的竖直距离之比,公式为 $R_{e,s} = d_e/d_s$, $d_e = |y_{e,l} - y_{e,r}|$,其中, $y_{e,l}$ 为左手肘对应y轴坐标, $y_{e,r}$ 为右手肘对应y轴坐标。

R_s 出现较大偏移时,表示人体姿态存在较大侧向倾斜,发生侧向跌倒可能性较大;偏移较小时身

体姿态较为对称,侧向跌倒的风险较小。肘部与肩部之间的相对偏移量 $R_{e,s}$ 较大时,表示身体重心向一侧失衡发生更大程度倾斜。

根据胯部-膝盖中心点在 y 轴上的竖直位置计算 $y_{k,h}$,并以相同方式计算肩膀-手腕 $y_{w,s}$,公式为 $y_{k,h}=y_{k,c}-y_{h,c}$, $y_{w,s}=y_{w,c}-y_{s,c}$ 。其中, $y_{k,c}$ 表示膝盖中心点 y 轴坐标, $y_{h,c}$ 表示胯部中心点 y 轴坐标, $y_{w,c}$ 表示手腕中心点 y 轴坐标, $y_{s,c}$ 表示肩膀中心点 y 轴坐标。

$R_{k,h}$ 、 $R_{w,s}$ 分别定义为大腿、手臂与人体外接矩形框高度的相对偏移程度,公式为 $R_{k,h}=|y_{k,h}|/h$, $R_{w,s}=|y_{w,s}|/h$ 。

与相对距离不同, $y_{k,h}$ 、 $y_{w,s}$ 反映大腿和手臂的高度关系。当发生前向跌倒时,手臂和膝盖常作为支撑点以减轻受伤程度,在后向跌倒时二者无法及时提供支撑,导致身体失去平衡,以此区分前向和后向跌倒。

方向判别得分计算如表3所示,其中, S_{fall} 表示侧向跌倒, F_{fall} 表示前向跌倒, B_{fall} 表示后向跌倒。在完成方向判别流程后,将得分最高的姿态作为结果输出。如果 $R_s>0.3$,表示身体侧向发生较大偏移,侧向跌倒得分增加0.7,进一步判断肘部与肩部在竖直方向上相对偏移量, $R_{e,s}>1$ 继续增加0.3;如果 $R_s\leq 0.3$,则前向和后向跌倒得分增加。膝盖-胯部高度差 $y_{k,h}>0$,表示膝盖低于胯部,前向跌倒得分增加0.5,根据相对偏移量 $R_{k,h}$ 进一步判断,当 $R_{k,h}>0.2$ 时,前向跌倒得分继续增加0.3; $y_{k,h}\leq 0$ 表示膝盖高于胯部,后向跌倒概率高,得分增加0.5,考虑到存在前向跌倒的可能性,得分增加0.2。手腕-肩膀高度差 $y_{w,s}>0$,表示手腕位于肩膀下方,前向跌倒得分增加0.5,进一步判断相对偏移量 $R_{w,s}$,当 $R_{w,s}>0.2$ 时,前向跌倒得分继续增加0.2;否则判定为后向跌倒,得分增加0.5。

表3 方向判别得分计算表

Table 3 Calculation table for direction discrimination score

姿态估计	判别条件	得分调整
肩膀偏移量 R_s	$R_s>0.3$	$S_{\text{fall}}+=0.7, R_{e,s}>1: S_{\text{fall}}+=0.3$
	$R_s\leq 0.3$	$F_{\text{fall}}+=0.3, B_{\text{fall}}+=0.5$
膝盖-胯部高度差 $y_{k,h}$	$y_{k,h}>0$	$F_{\text{fall}}+=0.5, R_{k,h}>0.2: F_{\text{fall}}+=0.3$
	$y_{k,h}\leq 0$	$B_{\text{fall}}+=0.5, F_{\text{fall}}+=0.2$
手腕-肩膀高度差 $y_{w,s}$	$y_{w,s}>0$	$F_{\text{fall}}+=0.5, R_{w,s}>0.2: F_{\text{fall}}+=0.2$
	$y_{w,s}\leq 0$	$B_{\text{fall}}+=0.5$

通过多维度分析制定的得分计算规则可以动态调整姿态得分,从而实现对复杂人体姿态变化更精确、细致的检测。相比于深度学习方法,这种基于判别规则进行得分调整的姿态估计方法在计算复杂度和资源消耗方面具备显著优势。深度学习方法通常需要大量数据进行训练,并且对计算资源的需求较高。而得分计算方法能在较低的计算负担下保持较高的准确率,能够有效避免延迟和高能耗的问题,可在资源受限的终端设备中实现良好的姿态识别性能。

3 边缘计算终端硬件部署与实现

3.1 硬件平台概述

在硬件平台上部署基于深度学习的网络模型,其难点集中于计算性能、推理延迟以及系统功耗3个方面。Jetson Nano是NVIDIA推出的面向嵌入式场景的低功耗边缘计算平台,相比传统嵌入式平台上使用CPU在处理卷积等运算时的效率短板,Jetson Nano采用并行架构,显著提升了模型前向传播的速度与资源利用率。CPU部分采用四核ARM Cortex-A57处理器,主频最高可达1.43 GHz,并配备2 MB二级缓存,为系统提供基础的计算能力。在图形处理方面,该平台集成了128核的NVIDIA

Maxwell架构GPU,支持CUDA并行计算,算力理论峰值可达每秒472亿次浮点运算,使其具备在本地并行运行多个神经网络模型的能力。

在内存配置上,平台配备4 GB 64 bit的LPDDR4低功耗内存,可满足深度学习模型权重加载、视频缓存和特征图中间计算结果存储等需求,针对本文需要连续视频处理的场景,可实现多帧视频数据的稳定缓存。64 bit的内存位宽能够有效降低CPU与GPU之间的数据延迟,保证推理全流程的端到端耗时符合需求,保证跌倒检测系统的实时性。

在软件层面,Jetson Nano运行Ubuntu操作系统,确保了操作的流畅性和用户界面的友好性。

算法部署基于系统镜像及TensorRT库函数实现。Jetson Nano B01预置了TensorRT 8.2.1的推理加速引擎。TensorRT能够通过层间融合、张量融合、精度校准及自动调优等图优化技术,降低模型的推理延迟和显存占用。在模型转换环节,TensorRT还支持将PyTorch、TensorFlow等框架导出的ONNX模型转换为高度优化的推理引擎,在损失极少精度的前提下,将模型参数量压缩至原先的一半甚至更低。

图5为TensorRT加速策略。如图5(a)所示,TensorRT实现了纵向层级合并技术,将卷积层、偏置层和激活函数顺序排列的网络层合并为单一的CBR层,使得整个操作仅占用单个CUDA核心。此外,TensorRT还支持横向层级整合,如图5(b)所示,将具有相同输入张量和结构的多个 1×1 CBR层合并为一个更大的 1×1 CBR层,同样仅消耗一个CUDA核心。除了层级合并,TensorRT还优化了concat层的处理,如图5(c)所示,将concat层的输入数据传输到后续层,从而减少了传输延迟。这些优化策略使TensorRT能够在不损失精度的前提下,大幅提升模型在Jetson Nano边缘设备上的执行效率。

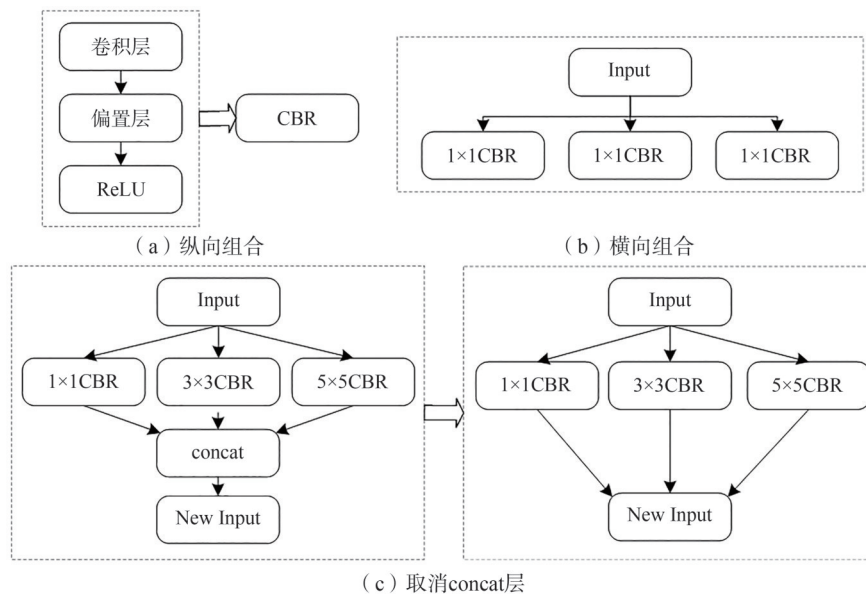


图5 TensorRT加速策略

Fig. 5 TensorRT acceleration strategy



图6 姿态识别权重文件转换流程

Fig. 6 Conversion process of pose recognition weight file

3.2 模型部署流程

算法模型在Jetson Nano上部署主要经过开发环境搭建、源码部署、模型量化、模型转换和Ten-

TensorRT推理优化5个步骤。为确保算子兼容性,从原始Pytorch训练框架中将权重文件先转换成ONNX序列化图,再生成适用于TensorRT推理的序列化引擎。权重文件转换流程如图6所示。

(1) 开发环境的搭建包括系统镜像与操作系统、深度学习框架和依赖库的安装。本研究在Ubuntu 20.04系统上进行操作,基于Jatpack 4.6.6系统镜像,配置conda-forge 3、python 3.8.10、OpenCV 4.2、CUDA 10.2、cuDNN 8.2.1.32、Pytorch 1.11、Torchvision 0.12为基础的虚拟环境。

(2) 虚拟环境搭建完成后将人体姿态检测模型和跌倒方向判别模型从服务器端拷贝到Jetson Nano上,包含权重文件、模型配置文件等。硬件终端联网后,配置虚拟环境中缺失的软件包以及OpenCV和Numpy等必要的依赖库。

(3) Pytorch模型量化随模型转换同时进行,首先通过重写机制将模型从浮点型格式转换为ONNX中间格式,再经过ONNX2trt将模型转换为TensorRT模型,精度由fp64降为fp32,最后通过8 bit量化,生成int8格式的TensorRT模型。

(4) ONNX模型转换完成后,使用TensorRT的Builder API进一步将模型转换为TensorRT引擎,在此过程中,TensorRT会进行层融合、精度校准和内存优化等操作,最后将构建好的引擎序列化为.engine文件,以便后续直接加载使用。

(5) 在TensorRT优化推理阶段,加载引擎并分配输入与输出内存后,系统会通过异步CUDA流管理数据传输与计算任务的重叠执行,从而最大化GPU利用率。同时创建执行上下文接口,将输入数据从CPU拷贝到GPU,执行推理后将输出数据从GPU拷贝回CPU,最终通过调整Jetson Nano的功耗模式和GPU频率,优化模型的推理速度和资源利用率。

完成以上流程后,TensorRT模型成功部署到Jetson Nano硬件终端。模型部署代码均采用python编程语言实现,通过调用python接口实现模型加载、运行环境初始化以及engine模型推理。

4 实验

4.1 自建数据集

选择COCO2017数据集为基础数据集,为贴合实验要求,基于本文跌倒姿态估计任务对数据集进行预处理,提取分类为person的图像和对应标签。

为提升模型在真实智慧养老场景中的鲁棒性,采用真实场景采集结合多维度增强的数据集构建策略。由于原COCO2017数据集中包含的人体姿态多为正常姿态,而老年人跌倒是日常生活中的小概率突发事件,真实环境下的跌倒数据难以采集,为提升模型对跌倒姿态的检测精度和泛化能力,需筛选日常活动场景下各类跌倒姿态图像构建数据集,包含室内、室外、各类运动场景下,不同年龄段人群的跌倒动作样本共1840张图像,覆盖老年人活动的各类场景及关键干扰因素,贴近实际应用环境。使用Roboflow软件进行人体边界框和17个骨骼关键点的标注。

为满足在光照多变、人体姿态易被遮挡等状况下精准检测的需求,提升模型对复杂场景的鲁棒性和泛化能力,针对跌倒姿态数据集采用以下方法进行数据增强。首先,对数据集进行水平翻转。然后,采用随机亮度变化,通过增加或减少像素亮度值,模拟日常环境下白天与晚上、晴天与阴天的不同光照条件。其次,增加HSV色域抖动动态调整亮度、饱和度参数,覆盖极端光照条件,解决传统单一对比度调整与实际光照变化差距大的问题。最后,添加高斯模糊和随机噪声,模拟实际环境拍摄过程中监控设备老化、拍摄距离变化导致的图像质量下降情况。部分数据集增强结果如图7所示。

通过数据集增强,筛掉不合理的图像并进行重制后,数据集总量扩充4倍,加上原COCO2017数据集,最终构建的数据集共包含70302张图像,将数据集命名为Fall_state,按照8:1:1划分为训练集、验证集、测试集,最终数据集包含56246张训练图像、7028张验证图像和7028张测试图像。



图7 部分数据集增强结果

Fig. 7 Selected dataset enhancement results

4.2 实验参数设置与评价指标

基于 PyTorch 开源深度学习框架完成模型训练与性能测试。在硬件层面,实验运行于搭载 Ubuntu 22.04 操作系统的服务器,CPU 选用 Intel (R) Platinum 8352V,主频 2.10 GHz,GPU 采用 NVIDIA RTX 4090 以支撑大批次训练需求。软件编译环境配置为 PyTorch 2.1.2、CUDA 12.1 和 Torchvision 0.16.2,编程语言采用 Python 3.8。

实验参数设置如下:图像像素尺寸为 640×640 ,批量大小设为 64,初始学习率为 0.01,权重衰减系数为 0.0005,最大训练轮次为 1000,早停轮数设置为 100,其他参数采用模型标准化配置,以保障训练过程的收敛性与稳定性。

以预测框与真实标注框的交并比(IoU)作为样本分类的判定基准。针对正样本场景,预测框与对应真实标注框的 IoU 值高于预设阈值,则该预测为正确检测,即真正例(True positive, TP)。若 IoU 值未达到阈值则为错误检测,即假正例(False positive, FP)。针对负样本场景,若某一预测框与全部真实标注框的 IoU 值均低于预设阈值,则该预测为真负例(True negative, TN),反之则为假负例(False negative, FN)。选取以下指标用于算法性能评估。

(1) 精确率(Precision, P_{rec}),用于衡量模型预测为正类的样本中实际为正类的占比,公式为 $P_{\text{rec}} = X_{\text{TP}} / (X_{\text{TP}} + X_{\text{FP}}) \times 100\%$ 。

(2) 召回率(Recall, R_{ec}),表征实际为正类的样本中被模型正确识别为正类的比例,用于评估模型的漏检情况,公式为 $R_{\text{ec}} = X_{\text{TP}} / (X_{\text{TP}} + X_{\text{FN}}) \times 100\%$ 。

(3) mAP 值,为所有类别的精度平均值,公式为 $\text{AP} = \int_0^1 P(r) dr$, $\text{mAP} = \frac{1}{c} \sum_{j=1}^c \text{AP}_j$, mAP@50 是 IoU 阈值为 0.5 时的精度平均值。

(4) 参数量(Parameters),用来表示模型的空间复杂度,单位为 M。

(5) GFLOPs(Giga floating-point operations per second),为每秒十亿次浮点运算次数,用来表示推理过程中的计算负载,数值越低,模型的推理效率越高,时间复杂度越低。

4.3 注意力机制性能对比实验

为了验证 SA 基于跌倒判别的优势,考虑到计算资源有限,在自建数据集上选择常见的轻量级注意力机制网络进行性能对比实验。基于跨空间学习的高效多尺度注意力(Efficient multi-scale attention, EMA)将通道和空间信息结合,采用多尺度并行子网络结构将通道维度分组为多个子特征,使空间语义特征在每个特征组内均匀分布,保留了通道维度的信息并降低了计算负担。无参数的注意

力机制(Simple attention module, SimAM)使用空间抑制理论设计能量函数,计算每个神经元的注意力重要性,通过计算得到的注意力重要性为每个神经元分配相应的注意力权重,从而提高特征图的表征能力。

表4为注意力机制网络模型的性能对比结果,实验结果表明,SA性能最好,参数量比EMA减少0.16 M,而精确率提高1.8个百分点,模型计算量减少了0.3 GFLOPs,模型复杂度更小。相比于SimAM,在参数量和计算复杂度相近的情况下精确率提高0.6个百分点。表明SA为基于骨骼关键点判别人体姿态和跌倒方向提供了精确率和效率的双重保障。

表4 注意力机制网络模型性能对比

Table 4 Performance comparison of attention mechanism networks

注意力机制	$P_{rec}/\%$	参数量/M	计算量/GFLOPs
EMA	84.0	2.78	7.5
SimAM	85.2	2.62	7.2
SA	85.8	2.62	7.2

4.4 消融实验

为研究所提算法对跌倒检测性能的影响,在数据集上进行消融实验和算法性能分析。表5是消融实验的具体设计方案,实验1为YOLOv11-Pose基础网络模型,实验2使用MobileNetV4轻量化网络UIB模块改进主干网络,实验3在原始网络基础上添加SA,实验4结合UIB模块和SA改进主干网络。实验结果显示,改进算法在参数量最少的情况下,Box和Pose的精确率分别为88.0%和84.8%(Box表示外接矩形框,Pose表示骨骼关键点),mAP@50分别为91.5%和81.2%。mAP@50对比基础模型均有轻微下降,分别降低0.4个百分点和0.7个百分点,但算法改进后参数量减少约14%,实现了计算量和精确率的平衡。

表5 消融实验结果对比

Table 5 Comparison of ablation experiment results

编号	YOLOv11-Pose	UIB	SA	参数量/M	Box			Pose		
					$P_{rec}/\%$	$R_{ec}/\%$	mAP@50/%	$P_{rec}/\%$	$R_{ec}/\%$	mAP@50/%
1	✓	×	×	2.87	88.1	83.8	91.9	85.0	77.0	81.9
2	✓	✓	×	2.73	88.3	83.7	92.0	85.3	76.7	81.6
3	✓	×	✓	2.62	88.2	82.8	91.4	85.2	76.3	81.3
4	✓	✓	✓	2.48	88.0	82.9	91.5	84.8	76.3	81.2

4.5 算法实际检测效果

对本文算法进行人体姿态识别和跌倒方向判别实验验证。图8为人体姿态检测效果图,分别对站立、坐姿、跌倒姿态进行可视化对比分析。图中蓝色框标注人体外接矩形框,将人体骨骼关键点进行连接,蓝色标签显示置信度,红色标签显示人体边界矩形框ID和目标检测结果。其中清晰图像的定位置信度均高于0.85,模型在图像画质较差或环境复杂的场景下也有良好的检测效果。结果表明,改进后的网络可以准确定位人体边界矩形框和骨骼关键点,所提算法能够正确识别人体姿态,区分日常行为和跌倒行为。

对识别为跌倒姿态的图像进一步进行跌倒方向判别,检测为前向跌倒输出跌倒状态“Ffall”,侧向跌倒输出跌倒状态“Sfall”,后向跌倒输出跌倒状态“Bfall”,结果如图9所示。实际场景中老年人跌倒图像难以大量采集,因此借助各类场景下跌倒图像进行算法验证。结果表明,所提算法能够正确

判别跌倒方向。

4.6 硬件终端检测效果

为了进一步验证系统的实际应用效果,结合 Jetson Nano 硬件设备、电容屏和 IMX219 摄像头搭建硬件平台。Jetson Nano 基于 Maxwell GPU 架构,集成了多项硬件优化模块。

ARM Cortex-A57 CPU 负责调度硬件系统的各项任务,MIPI CSI 接口读取存储在 SD 卡中的系统镜像和经过转换的模型权重文件,CPU 负责加载模型参数并将其传输至 GPU。同时,CPU 通过 MIPI CSI 接口获取来自 IMX219 摄像头采集的实时图像数据,由 GPU 执行目标检测与姿态识别的推理任务,生成检测结果,包括目标类别和姿态类别等信息。推理完成后,CPU 将 GPU 计算得到的目标检测与姿态识别结果传递至高清音视频子系统,进行图像合成和文字叠加等后处理操作。显示控制综合体 DCC 将最终的检测画面通过 HDMI 接口实时输出至显示设备,同时内存控制器 MC 通过 MicroSD 接口将检测结果存储至本地 SD 卡, Micro USB 接口将数据传输至终端服务器,实现对姿态识别任务的实时监控与数据管理。

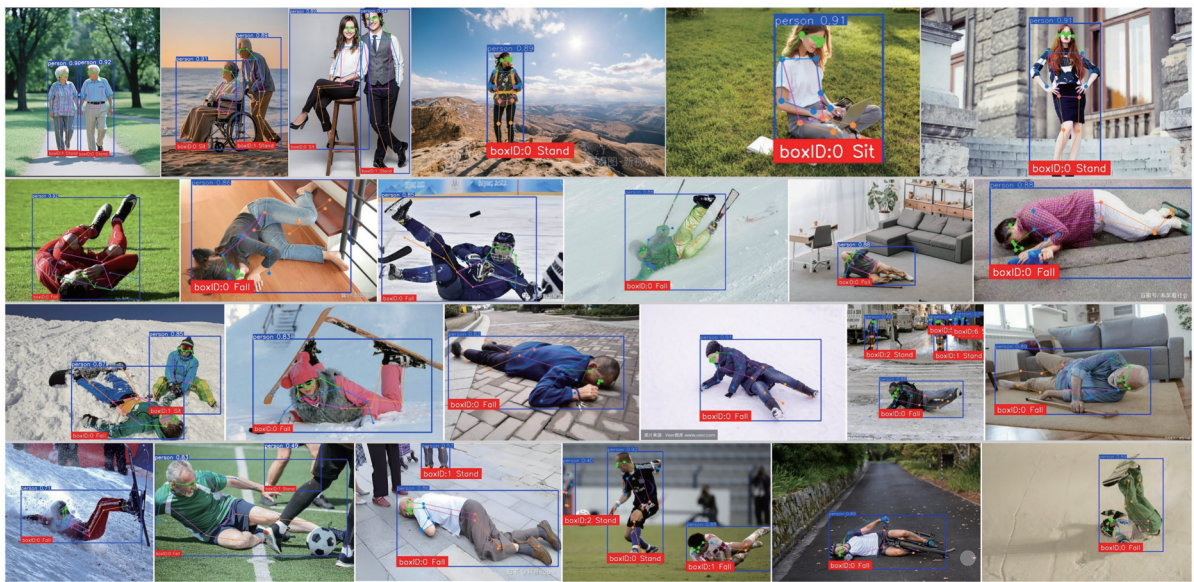


图8 人体姿态检测效果图
Fig. 8 Diagram of pose detection result



图9 跌倒方向判别效果图
Fig. 9 Diagram of fall direction discrimination result

将硬件终端布设于模拟居家养老的实验场景中,对突发跌倒行为的视频帧进行实时检测分析。检测流程及硬件终端部署如图 10 所示。

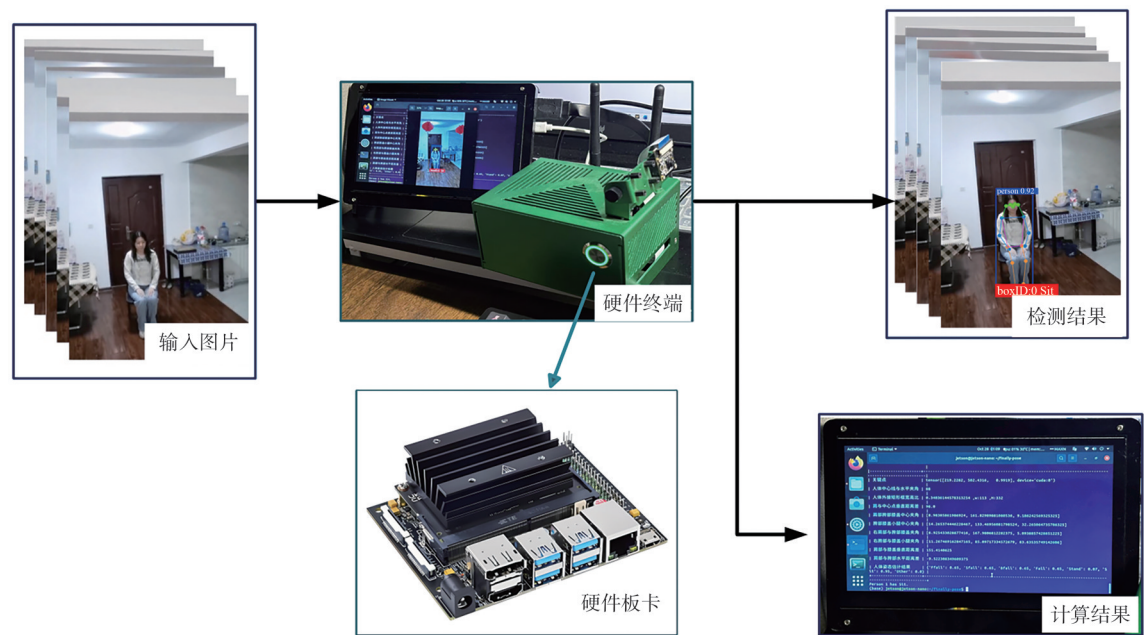


图 10 硬件终端部署
Fig. 10 Hardware terminal deployment

首先由采集模块获取模拟居家场景下的输入图像或视频流,将数据传输至集成 Jetson Nano 硬件终端。终端内部基于优化后的 TensorRT 引擎执行并行推理计算,依次完成人体目标检测、骨骼关键点提取与跌倒状态判别。推理完成后,系统同步输出两类结果,一类是进行了检测框和关键点标注,显示跌倒状态标签的可视化检测结果,另一类是推理耗时和得分计算细节表,实现了数据管理。

图 11 为算法在实际场景中的检测结果,所提算法能够准确识别图像中目标人体的全部骨骼点并连接人体骨架图,可以有效区分人体正常姿态与跌倒姿态,并针对不同的跌倒方向实现准确判别。



图 11 硬件终端姿态识别效果图
Fig. 11 Diagram of hardware terminal pose recognition effect

实验结果表明,改进算法在实际场景中具有较高的实用性和可靠性,边缘计算架构有效解决了传统跌倒检测中响应慢与隐私保护问题。同时,跌倒方向判别能够通过方向信息为智慧养老救治提供可靠依据,有效提升救治效率。

5 结论

本文以边缘计算为核心技术架构,通过跌倒检测算法优化、轻量化模型硬件部署与多特征融合算法设计,有效解决了传统跌倒检测技术信息传输响应慢以及隐私安全无法保障两大核心问题。采用 MobileNetV4 与 SA 优化算法,使其适配边缘算力,在硬件终端实现 88.0% 的人体外接矩形框检测准确率与 84.8% 的骨骼关键点检测准确率。同时,所提算法实现了跌倒方向判别,能够帮助医护人员提前明确不同跌倒方向对应的差异化受伤风险,优化资源配置流程,有效提升救治效率,适配智慧养老服务的发展需求。未来研究可进一步优化硬件终端与智慧通信的协同机制,一方面降低终端设备与通信网络的交互频率以减少资源消耗,另一方面探索多硬件终端协同计算模式,实现跌倒预警、急救调度、健康管理的闭环服务,构建更高效、更安全的智慧养老生态体系。

参考文献:

- [1] 张莉,吴泽兰. 数字背景下的养老服务体系策略创新[J]. 数字技术与应用,2022,40(10):39-41.
- [2] 张泽漓. 智慧赋能养老服务的驱动要素、转型逻辑、实践困境与对策[J]. 西安交通大学学报,2024,44(3):64-74.
- [3] 张丽艳,石皓月. 健康老龄化视角下中国智慧健康养老政策研究:基于2015—2020年政策文本分析[J]. 科学与管理,2022,42(3):27-33.
- [4] Gutiérrez J, Rodríguez V, Martín S. Comprehensive review of vision-based fall detection systems[J]. Sensors, 2021,21(3):947.
- [5] Alarifi A, Alwadain A. Killer heuristic optimized convolution neural network-based fall detection with wearable IoT sensor devices[J]. Measurement, 2021,167:108258.
- [6] Yazar A, Erdren F, Cetin A E. Multi-sensor ambient assisted living system for fall detection[C]//Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Florence:IEEE,2014:1-3.
- [7] Xu Z M, Ding J Y, Zhang S S, et al. mmCMD: Continuous motion detection from visualized radar micro-doppler signatures using visual object detection techniques[J]. IEEE Sensors Journal, 2024,24(3):3394-3405.
- [8] 米增,连哲. 面向通用目标检测的YOLO方法研究综述[J]. 计算机工程与应用,2024,60(21):38-54.
- [9] Min W D, Wei L S, Han Q, et al. Human fall detection based on motion tracking and shape aspect ratio[J]. International Journal of Multimedia and Ubiquitous Engineering, 2016,11(10):1-14.
- [10] Min W D, Cui H, Rao H, et al. Detection of human falls on furniture using scene analysis based on deep learning and activity characteristics[J]. IEEE Access, 2018,6:9324-9335.
- [11] Xu Q Z, Huang G Y, Yu M J, et al. Fall prediction based on key points of human bones[J]. Physica A: Statistical Mechanics and Its Applications, 2020,540:123205.
- [12] Chen W M, Jiang Z J, Guo H L, et al. Fall detection based on key points of human-skeleton using OpenPose[J]. Symmetry, 2020,12(5):744.
- [13] Lou H T, Duan X H, Guo J M, et al. DC-YOLOv8: Small-size object detection algorithm based on camera sensor[J]. Electronics, 2023,12(10):2323.

- [14] Qin D F, Leichner C, Delakis M, et al. MobileNetV4: Universal models for the mobile ecosystem[C]//Proceeding of the 18th European Conference on Computer Vision. Milan, Italy: Springer-Verlag, 2024: 78-96.
- [15] Yang Y B. SA-Net: Shuffle attention for deep convolutional neural networks[C]//2021 IEEE International Conference on Acoustics, Speech and Signal Processing. Toronto, ON, Canada: IEEE, 2021: 2235-2239.

(责任编辑 孙 锴)

Design and Implementation of an Edge Computing Terminal for Fall Detection in Smart Elderly Care

LIU Fangyu, FAN Yun, QI Ruian, HUANG Wei

(School of Electronic Information Engineering, Inner Mongolia University, Hohhot 010070, China)

Abstract: The deep integration of the digital economy and artificial intelligence has accelerated the progress of smart city construction. As a crucial supporting segment of people's livelihood services in smart cities, smart elderly care imposes higher requirements on information transmission within its scenarios. Aiming at the two core issues in current fall detection, including prominent delays in information transmission and response as well as significant privacy and security risks, a hardware system for fall posture classification is designed and implemented. The system adopts an edge computing architecture, with the improved YOLOv11-Pose as its core. It optimizes the backbone network using MobileNetV4 to reduce computational complexity, and simultaneously develops a multi-dimensional scoring-based algorithm integrating skeletal point features to realize fall behavior recognition and fall direction determination. The algorithm is deployed on the Jetson Nano hardware terminal, enabling real-time responses at the edge. This not only significantly reduces the amount of data transmission but also avoids the risk of privacy leakage. Experimental results demonstrate that the proposed method can accurately distinguish different postures and correctly determine fall directions. The test results of the edge hardware terminal show that the system has low discrimination latency and stable operation, providing efficient and reliable technical support for smart elderly care scenarios.

Key words: edge computing; human pose estimation; fall direction determination; YOLOv11-Pose