

激光与光电子学进展

面向嵌入式系统的人体行为识别

伏娜娜, 刘大铭*, 张恒博, 李讓洞

宁夏大学物理与电子电气工程学院, 宁夏 银川 750021

摘要 为了在嵌入式平台上使人体行为识别网络达到实时效果,提出了一种基于轻量级 OpenPose 模型的人体行为识别方法。所提方法从人体的 18 个骨骼关键点角度出发,通过骨骼关键点的空间位置确定行为类型。首先通过轻量级 OpenPose 模型提取人体的 18 个骨骼关键点坐标信息,然后利用关键点的编码对人体的行为进行描述,最后利用分类器对获取的关键点坐标进行分类,从而识别出人体的行为状态,并将其移植到 Jetson Xavier NX 设备上,利用单目相机进行了测试。实验结果表明,所提方法在嵌入式开发板 Jetson Xavier NX 上能够快速、准确识别出行走、挥手、下蹲等人体的 11 类行为,平均识别准确率达到 96.08%,检测速度达到了 11 frame/s 以上,相比于原模型,检测速度提升了 177%。

关键词 机器视觉; 嵌入式系统; Jetson Xavier NX; OpenPose; 行为识别; 轻量级网络

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.2215001

Human Behavior Recognition for Embedded System

Fu Nana, Liu Daming*, Zhang Hengbo, Li Xuandong

College of Physics, Electronics and Electrical Engineering, Ningxia University,
Yinchuan 750021, Ningxia, China

Abstract To achieve real-time effects of the human behavior recognition network on the embedded platform, a human behavior recognition technique based on the lightweight OpenPose model is proposed. This approach begins with the viewpoint of 18 human body bone key points and calculates the behavior type based on the spatial position of the bone key points. First, the lightweight OpenPose model is used to extract the 18 bone key points to coordinate information about the human body. Then, the key point coding is used to describe the human body behavior. Finally, the classifier is used to classify the acquired key point coordinates to detect the human body behavior status and transplant it into Jetson Xavier NX equipment using a monocular camera for testing. Experimental results show that this method can quickly and accurately identify 11 types of human behaviors, such as walking, waving, and squatting, on the embedded development board Jetson Xavier NX, with an average recognition accuracy rate of 96.08%, and detection speed of >11 frame/s. The frame rate is increased by 177% compared to the original model.

Key words machine vision; embedded system; Jetson Xavier NX; OpenPose; behavior recognition; lightweight network

1 引言

近几年,随着计算机视觉的迅速发展,人体行为识别技术也发展得如火如荼,并广泛应用于各个领域,如检测独居老人异常行为、体感游戏等。因此,人体行为识别得到越来越多相关领域人员的关注,并且成为计算机视觉领域的研究热点^[1]。

对人体行为识别的研究主要分为两个方面:一方面是基于静态图像的;另一方面是基于视频的^[2]。基

于静态图像的人体行为识别,主要通过单张静态图像提取人体行为特征,然后利用提取到的特征进行分类,然而这种方法存在很多的局限性,比如跑步是一种连续动态的人体行为,而只靠一张静态图像无法对跑步行为进行准确识别。基于视频的人体行为识别,主要将视频视为连续静态图像的时间序列。目前,基于视频的人体行为识别已经取得了很不错的成果,但是仍存在一些难点:1)有遮挡物体时,不能准确识别人体被遮挡的部分,并且识别速度较慢^[3];2)在密集人

收稿日期: 2021-08-10; 修回日期: 2021-09-02; 录用日期: 2021-09-24

基金项目: 宁夏重点研发计划重大项目(2018BBF02022-04)、宁夏自然科学基金(2021AAC03113)

通信作者: *nxldm@126.com

群中无法准确判断出图像中的人数,不能提取出有效的人体特征,导致行为识别的准确率不高^[4];3)深度学习的模型一般较大,将其部署到嵌入式等设备上,将面临设备计算能力不足的问题,导致行为识别的实时性不高;4)针对日常活动的人体行为检测,需要特定的数据集。

为了解决这些问题,本文利用 MobileNet 中深度可分离的思想对 OpenPose 模型的特征提取网络 VGG19 进行改进,提出了一种基于轻量级 OpenPose 模型的人体行为识别方法。首先拍摄并制作了一个包括日常活动中较频繁出现的 11 类行为的人体行为的数据集,这 11 类行为分别为站立、行走、跑步、挥手、弯腰、踢腿、拥抱、操作电脑、侧压腿、跌倒和下蹲。然后将模型部署在 Jetson Xavier NX 上,构建人体行为识别系统,实现了算法和应用的对接。Jetson Xavier NX 是一款功能强大的人工智能开发板,设备体积小性能高。该模块拥有超强的运算能力,在 15 W 的功率下可以达

到每秒 14 万亿次运算,它和信用卡一般大,并支持多个神经网络的并行运算^[5]。

2 系统结构

基于嵌入式 Jetson Xavier NX 平台的人体行为识别系统结构如图 1 所示,由 Jetson Xavier NX 开发板、显示屏、摄像头、榛子云平台 and 手机组成。摄像头采用海康威视的 USB 高清摄像头,并连接到 Jetson Xavier NX 的 USB 接口上,用来拍摄人体日常活动中的行为图像;显示屏与 Jetson Xavier NX 的 HDMI 接口相连,用来显示摄像头拍摄的画面、当前画面中人的数量、当前识别的人体行为的名称、检测速度大小及检测到跌倒行为时的警告信息等。榛子云与 Jetson Xavier NX 远程连接,当系统识别到跌倒行为时,系统通过榛子云平台及时将识别结果反馈给手机。Jetson Xavier NX 作为主控制器,将训练好的人体行为识别模型移植到其中,构成人体行为识别主控设备。

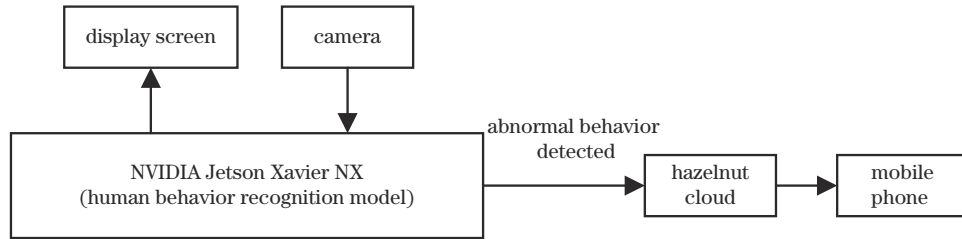


图 1 系统结构框图

Fig. 1 Block diagram of system structure

3 轻量级 OpenPose 模型

3.1 OpenPose 模型

本实验组选择了目前主流的多人姿态估计,多人姿态估计主要有两大类^[6]:一类是运行效率较低、检测精度高的自顶向下的方法,典型的模型有 RMPE^[7]等;另一类是运行效率高、但检测精度略低的自底向上的方法,最典型的模型是 OpenPose^[8]。考虑到嵌入式平台 Jetson Xavier NX 计算能力的问题,为了保证检测过程的实时性,选择 OpenPose 模型来提取人体的骨骼关键点,OpenPose 模型的网络结构如图 2 所示。

输入原始图像,经过 VGG19^[9]前 10 层后生成特征图 F ,然后将 F 作为后面卷积神经网络的输入。在 OpenPose 中,以卷积块 C(图 2 中 3 个连续的 3×3 卷积)代替了原始方法中 7×7 卷积。人体姿态估计包括多个阶段双分支的卷积神经网络,在每个阶段,上支链均用来预测关节热度图 S' ,下支链均用来预测关节亲和域 L' ,其中 ρ^1 和 ϕ^1 为第 1 阶段用于推理的卷积网络,分别对 F 进行处理,生成关节热度图 $S^1 = \rho^1(F)$ 和关节亲和域 $L^1 = \phi^1(F)$ 。 ρ^l 和 ϕ^l 为第 l 阶段用于推理的卷积网络,每个阶段的输入为 F 和前一阶段的预测结

果,然后使用本阶段的卷积网络分别预测出本阶段的关节热度图和关节亲和域:

$$S' = \rho^l(F, S'^{-1}, L'^{-1}), \forall t \geq 2, \quad (1)$$

$$L' = \phi^l(F, S'^{-1}, L'^{-1}). \quad (2)$$

3.2 MobileNet

Google 公司于 2017 年 4 月提出 MobileNet^[10],该网络的基本单元是深度可分离卷积。深度可分离卷积将标准卷积分解为一个深度卷积和一个 1×1 的点卷积,这种分两步进行卷积的方式可以极大地减小网络参数量^[11]。具体分解过程如图 3 所示,其中[图 3(a)]是标准卷积,[图 3(b)]是深度卷积,[图 3(c)]是 1×1 的点卷积。

当输入特征图大小为 $D_F \times D_F \times M$ 、输出特征图大小为 $D_F \times D_F \times N$ 、卷积核大小为 $D_K \times D_K$ 时,标准卷积的计算量 C_S 为

$$C_S = D_K \times D_K \times M \times N \times D_F \times D_F, \quad (3)$$

深度可分离卷积的计算量 C_D 为

$$C_D = D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F, \quad (4)$$

二者计算量之比 P_C 为

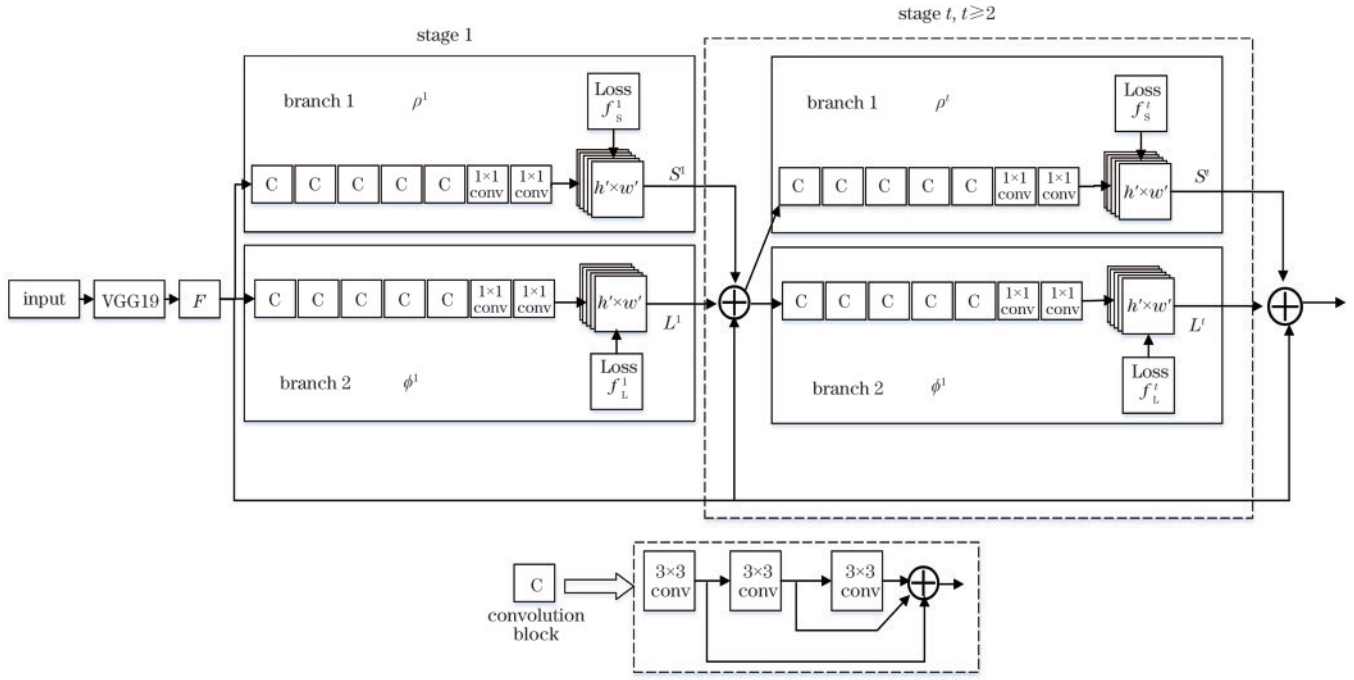


图 2 OpenPose 网络结构
Fig. 2 Network structure of OpenPose

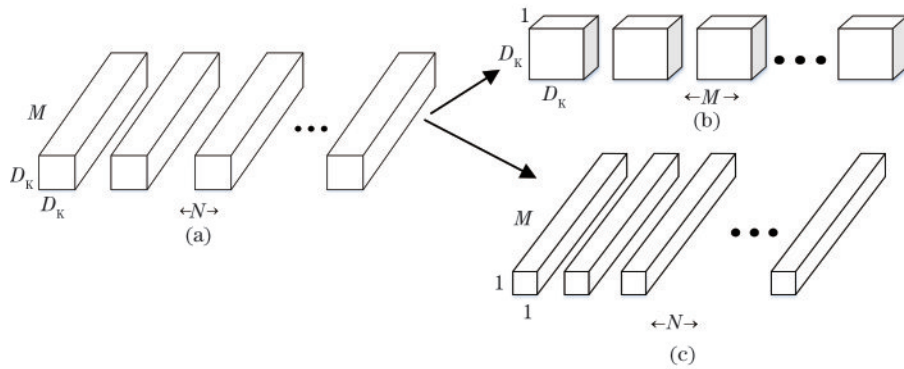


图 3 深度可分离卷积分解过程。(a)标准卷积;(b)深度卷积;(c)点卷积
Fig. 3 Depthwise separable convolution decomposition process. (a) Standard convolution; (b) depthwise convolution; (c) pointwise convolution

$$P_c = \frac{C_D}{C_s} = \frac{1}{N} + \frac{1}{D_k^2}, \quad (5)$$

式中： N 一般取较大的值，因此 $\frac{1}{N}$ 的结果很小。而 OpenPose 网络结构中的卷积核为 3×3 ，即 $D_k = 3$ ，由式(5)可得，深度可分离卷积的计算量仅约为传统卷积的 $1/9$ 。

3.3 基于 MobileNet 改进的 OpenPose 模型

由于 OpenPose 模型过大、网络参数量较多，不易部署在嵌入式平台，本实验组用轻量级网络 MobileNet^[10] 中的深度可分离卷积结构替换了原 OpenPose 模型中特征提取网络 VGG19 中的常规卷积结构。轻量级 OpenPose 模型的特征提取网络变为由 MobileNet^[10] 调整后得到的 12 层网络，其中第 1 层是一个传统卷积(conv)，其他 11 层均建立在深度可分离

卷积(conv dw)上，conv dw_7 层中的深度卷积扩张系数(dilation)设置为 2，填充(padding)参数也为 2，其他层的扩张系数均为 1，不填充，所有卷积核的大小为 3×3 ，并且卷积核的数量随着层数的加深而增多。调整后的 12 层特征提取网络结构如表 1 所示。

3.4 人体行为识别模型

在姿态估计获得的 18 个人体骨骼关键点的基础上，利用关键点的坐标信息进行人体行为识别。人体行为识别的研究并不需要提取人体的全部特征，只需要通过主要关节空间位置来描述，因此，只要 18 个关键点的组合便能形成诸多类行为。本实验组采集了大量不同行为的数据，利用不同行为之间关键点的空间位置差异，对每类行为每一帧的 18 个关键点的坐标进行标注，训练了 11 类行为：站立、行走、跑步、挥手、拥抱、踢腿、侧压腿、操作电脑、弯腰、下蹲及跌倒。将

表 1 调整后的特征提取网络结构

Table 1 Adjusted feature extraction network structure

Convolution type	Convolution kernel size	Stride	Dilation	Padding
conv	3×3×32	2	1	0
conv dw_1	3×3×64	1	1	0
conv dw_2	3×3×128	2	1	0
conv dw_3	3×3×128	1	1	0
conv dw_4	3×3×256	2	1	0
conv dw_5	3×3×256	1	1	0
conv dw_6	3×3×512	1	1	0
conv dw_7	3×3×512	1	2	2
conv dw_8	3×3×512	1	1	0
conv dw_9	3×3×512	1	1	0
conv dw_10	3×3×512	1	1	0
conv dw_11	3×3×512	1	1	0

训练好的人体行为识别模型移植到嵌入式端中,实现嵌入式平台的人体行为识别。

4 实验结果与分析

4.1 数据集

使用 OpenPose 姿态估计算法获取人体的 18 个骨骼关节点(0:鼻子;1:颈部;2:右肩;5:左肩等),各类行为均由这 18 个关节点组合形成,由这 18 个骨骼关节点构成的人体骨架图如图 4 所示。

收集了日常活动中的 11 类行为,包括站立、行走、跑步、下蹲、挥手、弯腰、拥抱、踢腿、跌倒、操作电脑和侧压腿,每类行为采集不同时长的视频,从采集的视频中提取出图像帧,共有 17454 张,利用这些图像标注的 18 个关键点制作成数据集,并将数据集按照 8:2 的比例随机划分为训练集和验证集,其中训练集为 13963 张,验证集为 3491 张,训练形成人体行为识别模型。该数据集中各类行为的样本数量如表 2 所示,各类行为的人体骨架图示例如图 5 所示。

4.2 网络训练

实验所使用的软硬件如表 3 所示。用自制的行为数据集进行人体行为识别模型的训练,其中训练集为

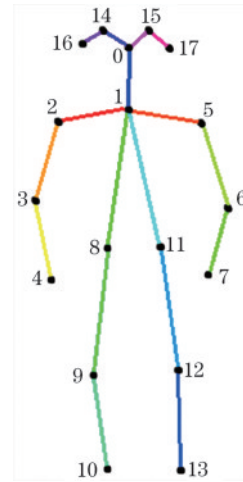


图 4 骨架图

Fig. 4 Skeleton diagram

表 2 数据集中各类行为样本数

Table 2 Number of samples of various behaviors in dataset

unit: frame

Category	11 types of human behavior data set (17454)		
	Number of samples	Number of samples in training set	Number of samples in validation set
Stand	1644		
Squat	1288		
Run	1608		
Bend	1428		
Fall	1008		
Operate the PC	2001	13963	3491
Leg press	2060		
Walk	1389		
Wave	782		
Kick	2300		
Hug	1946		

13963 张图像标注的关键点坐标,验证集为 3491 张图像标注的关键点坐标,标注格式为 csv,实验相关参数设置如表 4 所示。

表 3 实验使用的软硬件

Table 3 Software and hardware used in experiment

Software and hardware platform	Parameter
Embedded development board	NVIDIA Jetson Xavier NX
Operating system	Ubuntu 18.04
Deep learning framework	Tensorflow
CPU	6-core NVIDIA Carmel ARM®v8. 2 64-bit CPU
GPU	NVIDIA Volta™ Architecture 384 NVIDIA® CUDA® cores and 48 Tensor cores
CUDA	10.2
cuDNN	8.0
Programming language	Python 3.6

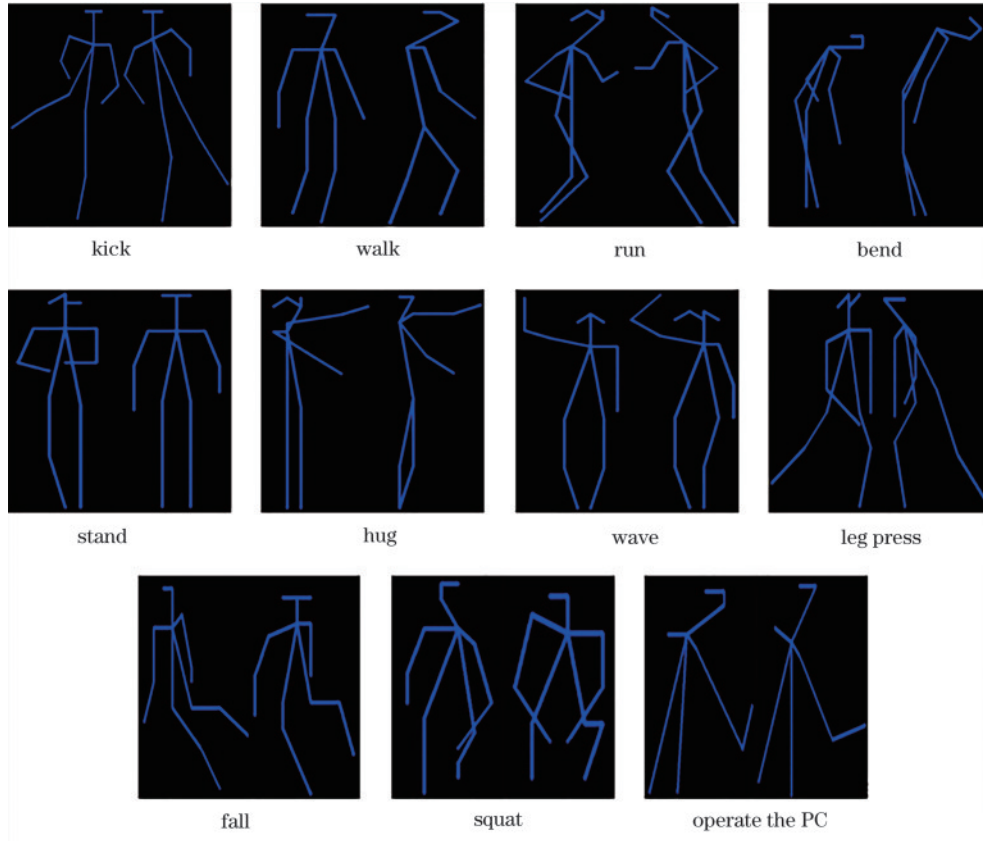


图 5 各类行为对应的骨架图示例图

Fig. 5 Examples of skeleton diagrams corresponding to various behaviors

表 4 实验参数说明

Table 4 Experimental parameter description

Parameter name	Parameter value
Input_size	224×224
Epoch	160
Batch_size	64
Learning_rate	0.0001
Loss function	Cross entropy loss
Optimizer	Adam

训练和验证过程的损失曲线和准确率 (acc) 曲线如图 6 所示, 其中横轴 Epoch 是数据集训练次数, 纵轴 Acc-loss 表示准确率和损失函数这两个指标, 右上角的 train acc 和 train loss 分别是训练准确率和训练损失, val acc 和 val loss 分别是验证准确率和验证损失。准确率指的是模型预测输出正确的行为分类数量占总输出行为分类数量的比例, 其表达式为

$$A_{acc} = \frac{\sum_{i=1}^n f(x_i) = y_i}{n}, \quad (6)$$

式中: i 为样本序号; x_i 为第 i 个行为数据样本; n 为样本总数; $f(x_i)$ 为模型预测输出; y_i 为真实标签。

采用交叉熵损失函数计算模型的损失函数值, 其表达式为

$$L_c = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log p_{ic}, \quad (7)$$

式中: N 为样本数量; i 为样本; c 为行为类别; M 为行为类别数量; y_{ic} 为符号函数 (0 或 1), 如果样本 i 的真实行为类别等于 c , 则 y_{ic} 取 1, 否则 y_{ic} 取 0; p_{ic} 为样本 i 属于行为类别 c 的预测概率。

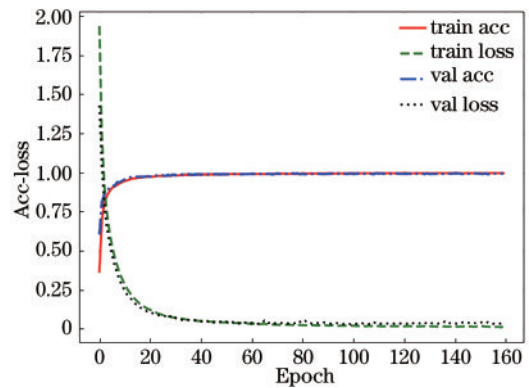


图 6 训练结果图

Fig. 6 Training result graph

从图 6 可以看出, 随着数据集训练次数的增多, 训练集与验证集的准确率曲线逐渐平稳, 并趋近于 1, 同时训练集与验证集的损失函数曲线也逐渐收敛, 并逐渐趋近于 0。在 Epoch 为 160 时, 训练集的准确率为 0.9947, 验证准确率为 0.9951, 训练损失为 0.0139, 验

证损失为 0.0295。而且在 Epoch 达到 40 以后,训练集与验证集的准确率曲线基本重合,说明人体行为识别模型训练效果较好。

4.3 Jetson Xavier NX 上的结果与分析

[图 7(a)]为实验测试装置,主要包括嵌入式开发板 NVIDIA Jetson Xavier NX、USB 摄像头、显示屏、显示屏支架、键盘等。[图 7(b)]为 Jetson Xavier NX,该开发板尺寸为 103 mm×90.5 mm×34 mm,显示输出有 HDMI 和 DP,USB 接口有 USB 3.1(4×),USB

2.0 Micro-B(1×)等。

为了验证基于 Jetson Xavier NX 人体行为识别的识别效果,于 2021 年 7 月在宁夏大学贺兰山校区 701 实验室开展了实验测试,利用单目 RGB 摄像头进行测试,测试者是同一个实验室的同学,为了更好地拍摄到整个人的活动范围,将摄像头固定在距地面 1.6 m 处,摄像头俯视角度设置为 30°。11 类人体行为识别混淆矩阵如表 5 所示,11 类人体行为测试成功的效果图如图 8 所示。



图 7 人体行为识别测试装置。(a)测试装置;(b)Jetson Xavier NX 开发板

Fig. 7 Human behavior recognition test device. (a) Test device; (b) Jetson Xavier NX development board

表 5 11 类人体行为识别混淆矩阵

Table 5 Recognition confusion matrix of 11 types of human behavior

unit: %

Category	Stand	Squat	Run	Bend	Fall	Operate the PC	Leg press	Walk	Wave	Kick	Hug
Stand	100	0	0	0	0	0	0	0	0	0	0
Squat	0	100	0	0	0	0	0	0	0	0	0
Run	1.3	0	94.5	0	0	0	0	4.2	0	0	0
Bend	0	16.7	0	83.3	0	0	0	0	0	0	0
Fall	0	2.5	0	0	97.5	0	0	0	0	0	0
Operate the PC	0	0	0	0	0	100	0	0	0	0	0
Leg press	0	3.3	0	0	5.7	0	91.0	0	0	0	0
Walk	1.5	0	1.2	0	0	0	0	97.3	0	0	0
Wave	0	0	0	0	0	0	0	0	100	0	0
Kick	0	0	0	0	0	0	0	0	0	100	0
Hug	0	0	0	0	0	0	0	0	6.7	0	93.3

表 5 给出了人体行为识别模型在嵌入式平台上测试 11 种不同人体行为的识别混淆矩阵,矩阵对角线上的数值表示行为识别的准确率,识别准确率的计算方法如下:只计算输出的正确行为为分类数量占总输出分类数量的比例,而不计算漏检情况^[12]。从表 5 可以看出,所提模型对站立、下蹲、操作电脑、挥手、踢腿行为的识别准确率较高,达到了 100%,而对弯腰行为的识别准确率较低,为 83.3%。通过分析可知,对于站立行为而言,18 个骨骼关键点容易被检测到,不易被遮挡,因此识别准确率较高;而弯腰、跌倒及侧压腿等行为的某几个骨骼点在进行行为时容易被遮挡,不易被检测,导致准确率较低。整体上来说,11 类行为的平均识别准确率为 96.08%,识别效果较为理想。

从图 8 可以看出,11 类行为均能很好检测到人体的骨骼关键点,检测速度为 9~11 frame/s,最高速度达到了 11.04 frame/s,这样的检测速度可以保证嵌入式平台上识别过程中的实时性。

为了验证轻量级 OpenPose 模型的有效性,将其与 OpenPose 模型进行了对比,以相同的方法在嵌入式平台上进行测试。用平均识别准确率和检测速度这两个指标来评价两种模型的性能,结果如表 6 所示。同时,为了验证所提方法的优越性,以识别行为类型和识别率为对比指标,将文献[4, 13-15]中的方法与所提方法进行了比较,结果如表 7 所示。

从表 6 可以看出,轻量级 OpenPose 在保证识别准确率损失很小的情况下,检测速度从 3.98 frame/s



图 8 测试成功的效果图

Fig. 8 Effect pictures of successful test

表 6 不同模型对比

Table 6 Comparison of different models

Model	Feature extraction network	File size /MB	Recognition accuracy /%	Detection speed / (frame · s ⁻¹)
OpenPose	VGG19	200	96.24	3.98
Lightweight OpenPose	MobileNet	7.5	96.08	11.04

提高到了 11.04 frame/s,提高了 7.06 frame/s,即检测速度提升了 177%。因此,轻量级 OpenPose 模型有效减小了计算量,提高了检测速度,更适用于嵌入式平台。

表 7 中,将所提方法与其他研究方法进行了比较,

所提方法不仅在识别行为类型上增加了几种日常生活中较频繁发生的类型,而且使行为识别准确率得到进一步提升。实验结果表明,所提模型在嵌入式平台上能够保证行为识别的实时性,也提升了人体行为识别算法的识别效果。

表 7 相关研究对比

Table 7 Related research comparison

Method	Type of behavior	Recognition rate /%
Reference[4]	clap, walk, dribble, play golf	86.25
Reference[13]	walk, jog, go up and down, sit, stand	91.60
Reference[14]	walk, run, go up and down, stand still, sit-stand, stand-sit, stand-squat, squat-stand	95.05
Reference[15]	walk, run, jump, go up and down stairs	85.00
Proposed method	stand, walk, run, squat, bend, kick, hug, fall, wave, side press, computer the PC	96.08

5 结 论

针对嵌入式平台上人体行为识别网络的实时效

果,提出了一种基于轻量级 OpenPose 模型的人体行为识别方法。首先通过轻量级 OpenPose 模型提取人体的 18 个骨骼关节坐标信息,然后利用关节的编码

对人体的行为进行描述,最后利用分类器对获取的关节点坐标进行分类。实验结果表明,该方法在嵌入式开发板 Jetson Xavier NX 上能够快速、准确识别行走、弯腰、下蹲、站立、挥手、跑步、踢腿、操作电脑、跌倒、拥抱、侧压腿等 11 类行为,平均识别准确率达到 96.08%,检测速度达到了 11 frame/s 以上,相比于原模型,检测速度提升了 177%。但是,该方法也存在一些局限性:1)人体行为的数据集偏小,且场景比较单一;2)嵌入式平台上识别的速度最高达到 11.04 frame/s,虽然能够满足实时性,但还是偏慢。针对上述问题,对未来工作的展望如下:1)在保证精度的情况下,进一步扩大数据集,尤其是复杂场景下的情况;2)进一步优化模型,以提升检测时的速度;3)将嵌入式开发板安装在小车底盘上或者机器人上,通过激光雷达和目标跟踪算法,移动实时捕捉人体的行为,更好实现人体行为识别。

参 考 文 献

- [1] 郭伏正,孔军,蒋敏. 自适应融合 RGB 和骨骼特征的行为识别[J]. 激光与光电子学进展, 2020, 57(20): 201506.
Guo F Z, Kong J, Jiang M. Action recognition based on adaptive fusion of RGB and skeleton features[J]. Laser & Optoelectronics Progress, 2020, 57(20): 201506.
- [2] 刘锁兰,顾嘉晖,王洪元,等. 基于关联分区和 ST-GCN 的人体行为识别[J]. 计算机工程与应用, 2021, 57(13): 168-175.
Liu S L, Gu J H, Wang H Y, et al. Human behavior recognition based on associative partition and ST-GCN [J]. Computer Engineering and Applications, 2021, 57(13): 168-175.
- [3] 任国印,吕晓琪,李宇豪. 基于 2D 转 3D 骨架的多特征融合实时动作识别[J]. 激光与光电子学进展, 2021, 58(24): 2410010.
Ren G Y, Lü X Q, Li Y H. Multi-feature fusion real-time action recognition based on 2D to 3D skeleton[J]. Laser & Optoelectronics Progress, 2021, 58(24): 2410010.
- [4] 李梦荷,许宏吉,石磊鑫,等. 基于骨骼关键点检测的多人行为识别[J]. 计算机科学, 2021, 48(4): 138-143.
Li M H, Xu H J, Shi L X, et al. Multi-person activity recognition based on bone keypoints detection[J]. Computer Science, 2021, 48(4): 138-143.
- [5] 唐乾琛. 英伟达公司发布全球最小边缘 AI 超级计算模块[J]. 科技中国, 2019(12): 108.
Tang Q C. NVIDIA releases the world's smallest edge AI supercomputing module[J]. China Scitechnology Business, 2019(12): 108.
- [6] 苏超,王国中. 基于改进 OpenPose 的学生行为识别研究[J]. 计算机应用研究, 2021, 38(10): 3183-3188.
Su C, Wang G Z. Research on student behavior recognition based on improved OpenPose[J]. Application Research of Computers, 2021, 38(10): 3183-3188.
- [7] Fang H S, Xie S Q, Tai Y W, et al. RMPE: regional multi-person pose estimation[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2353-2362.
- [8] Cao Z, Simon T, Wei S H, et al. Realtime multi-person 2D pose estimation using part affinity fields[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1302-1310.
- [9] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2021-05-03]. <https://arxiv.org/abs/1409.1556>.
- [10] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J]. International Journal of Computer Vision, 2017, 5(8): 122-131.
- [11] 胡佳玲,施一萍,谢思雅,等. 基于 Jetson nano 的改进 MobileNet 人脸识别系统[J]. 传感器与微系统, 2021, 40(3): 102-105.
Hu J L, Shi Y P, Xie S Y, et al. Improved Mobile Net face recognition system based on Jetson nano[J]. Transducer and Microsystem Technologies, 2021, 40(3): 102-105.
- [12] 苏寒松,刘腾腾,刘高华,等. 基于神经网络的学生行为检测算法研究[J]. 激光与光电子学进展, 2020, 57(22): 221016.
Su H S, Liu T T, Liu G H, et al. Algorithm for student behavior detection based on neural network[J]. Laser & Optoelectronics Progress, 2020, 57(22): 221016.
- [13] Catal C, Tufekci S, Pirmit E, et al. On the use of ensemble of classifiers for accelerometer-based activity recognition[J]. Applied Soft Computing, 2015, 37: 1018-1022.
- [14] 贾小云,王二虎,吴敬一. Android 平台下的实时人体行为识别[J]. 计算机工程与应用, 2018, 54(24): 164-167, 175.
Jia X Y, Wang E H, Wu J Y. Real-time human behavior recognition based on Android platform[J]. Computer Engineering and Applications, 2018, 54(24): 164-167, 175.
- [15] 刘宇,江宏毅,王仕亮,等. 基于加速度时域特征的实时人体行为模式识别[J]. 上海交通大学学报, 2015, 49(2): 169-172.
Liu Y, Jiang H Y, Wang S L, et al. Real-time human activity pattern recognition based on time domain features of acceleration[J]. Journal of Shanghai Jiao Tong University, 2015, 49(2): 169-172.