

# 基于 YOLOv5 算法的人体跌倒检测系统设计

周洪成, 杨娟, 徐志国

(金陵科技学院电子信息工程学院, 江苏 南京 211169)

**摘要:**针对老人、儿童、残障人士群体存在的跌倒风险,提出了一种基于 YOLOv5 算法的视频图像人体跌倒检测方法。该方法通过 CSI 摄像头采集视频数据,对视频数据进行训练和验证,进而判断目标人物的动作姿态,连接树莓派的 WiFi 模块将动作姿态信息发送到监护人手机上,提醒监护人对跌倒人员进行及时救治,从而提高了救援效率。实验结果表明:1)YOLOv5 算法对站立和跌倒动作的识别精确度均较高,而对下蹲动作的识别精确度相对较低;2)光线充足的环境中图片的置信度要高于昏暗环境中图片的置信度;3)YOLOv5 算法检测人体动作的速率和精确度高于 Faster R-CNN 算法。

**关键词:**人体跌倒;检测系统;树莓派;YOLOv5 算法

**中图分类号:**TP183

**文献标识码:**A

**文章编号:**1672-755X(2022)02-0022-08

## Design of Human Fall Detection System Based on YOLOv5 Algorithm

ZHOU Hong-cheng, YANG Juan, XU Zhi-guo

(Jinling Institute of Technology, Nanjing 211169, China)

**Abstract:** Aiming at the fall risk of the elderly, children and the disabled, a video image human fall detection method based on YOLOv5 algorithm is proposed. This method collects video data through the CSI camera, trains and verifies the video data, and then judges the action posture of the target person. The YOLOv5 algorithm connects the WiFi module of Raspberry Pi to send the action posture information to the guardian's mobile phone, and reminds the guardian to rescue the fall person timely. Thus YOLOv5 algorithm improves the rescue efficiency. The experimental results show that: 1) YOLOv5 algorithm has high recognition accuracy for standing and falling movements, but has relatively low recognition accuracy for squatting movements. 2) The confidence coefficient of pictures in sufficient light environment is higher than that of pictures in dark environment. 3) The speed and accuracy of YOLOv5 algorithm in detecting human motion are higher than that of Faster R-CNN algorithm.

**Key words:** human fall; detecting system; Raspberry Pi; YOLOv5 algorithm

随着人们生活水平的不断提高,人身安全越来越受到重视,安防系统也愈发受到人们的欢迎。世界卫生组织的一篇报道记载,跌倒造成的死亡率仍居高不下,全球每年有 30 多万人死于跌倒,其中一半是 60 岁以上的老人<sup>[1]</sup>。老年人、幼儿、残疾人等发生跌倒行为,会引发危险。

为了应对这些危险,需要在传统监控安防系统的基础上研制一种新型的监控安防系统,及时检测出跌

**收稿日期:**2022-05-14

**基金项目:**教育部产学研合作项目(202101352041);江苏省产学研合作项目(BY2021381);江苏省高等教育学会“十四五”研究规划项目(YB101);2021 年“大人网云”虚拟班建设项目(D2021005);金陵科技学院教改课题项目(JYJG202107)

**作者简介:**周洪成(1978—),男,江苏扬州人,副教授,博士,主要从事图像处理研究。

倒等危险行为,并发出警报预警,提醒用户对受伤人员采取有效的救治措施,从而保证人身安全。根据跌倒信号获取的装置不同,跌倒检测可分为三种:基于可穿戴设备的跌倒检测系统、基于环境信号的跌倒检测系统和基于视频监控的跌倒检测系统<sup>[2]</sup>。

Guo 等<sup>[3]</sup>将网络连接到家庭自动化系统,从而实现了无线网络通信的功能。Qiu 等<sup>[4]</sup>使用基于环境信号的跌倒检测系统,在实验场地中放置多个传感器,根据传感器传输的数据来判断目标的状态以及家庭安全情况。Chen 等<sup>[5]</sup>以视频监控、人工智能检测和入侵报警为切入点,设计了一套基于图像处理的多功能远程监控家庭智能安防系统。Islam 等<sup>[6]</sup>和徐毅等<sup>[7]</sup>利用物联网技术,通过远程控制智能家居和远程监控,实现全方位的安全检测。综上所述,已有成果对于连续视频检测人体跌倒方面的研究并不多。为此,本文提出了一种基于 YOLOv5 算法的视频图像人体跌倒检测方法,该方法在检测到跌倒时安防系统会发出警报,提醒监护人对跌倒人员进行救治。

## 1 YOLOv5 算法描述

目前,从图像中检测人物的方法大致分为两类:一类是提取人体运动学的显性知识作为行为判断的依据,例如:宽高比、重心等。这类方式存在很大的局限性,在分辨率低、光线过明或过暗、摄像头位置变动的情况下,都可能辨别不清人物的身体特征,从而不能实现人物跟踪<sup>[8]</sup>。另一类是基于深度学习的目标检测算法,如 YOLOv5 算法等。此算法使用深度学习技术,从图像中提取特征,根据训练好的模型来判断图像中是否有待检测目标。目标检测的结构一般分为两种(如图 1 所示):一种是 second stage,另一种是 first stage。它们的区别在于 second stage 具有区域方案过程,此过程类似于筛选过程,神经网络会根据候选区域提前生成位置和类别;first stage 则是直接生成位置和类别<sup>[9]</sup>。

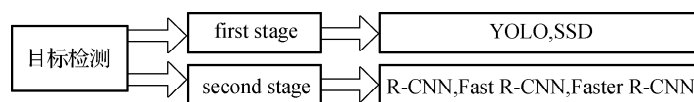


图 1 基于深度学习的目标检测算法

基于深度学习的 YOLOv5 算法是一种使用卷积神经网络(convolutional neural networks,CNN)特征的学习方法,这种学习方法能够自动发现、检测及分类目标所需的特征,并通过卷积神经网络将原始的输入信息转化成更抽象、更高维的特征,从而满足实际生产生活中的大多数应用需求。

在检测出目标后,还需要一些参数来评价检测的效果<sup>[10]</sup>。其中 TP(true positive)指被判定为正样本,事实上也是正样本;TN(true negative)指被判定为负样本,事实上也是负样本;FN(false negative)指被判定为负样本,事实上是正样本;FP(false positive)指被判定为正样本,事实上是负样本;AP(average precision)表示平均精确度;mAP(mean average precision)表示各个类别 AP 的平均值;精确度(precision,P)表示样本中正例被预测正确的目标数占有所有预测为正例的比率,该值越接近于 1,表明精确度越高<sup>[11]</sup>,公式如下:

$$P = \frac{TP}{TP + FP} \quad (1)$$

召回率(recall,R)表示预测样本中正例被预测正确样本数占有所有正例样本数的比例,公式如下:

$$R = \frac{TP}{TP + FN} \quad (2)$$

准确率(accuracy,ACC)表示预测样本中预测正确的样本数占有所有样本数的比例,公式如下:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

YOLOv5 算法的网络结构主要包括 4 个部分:输入端、Backbone、Neck、Prediction。输入端通过 Mosaic 进行数据增强;Backbone 主要包含 Focus 切片处理、跨阶段对等(cross stage parity,CSP)结构,起到切片下采样的功能;Neck 采用了特征金字塔网络(feature pyramid networks,FPN)+路径聚合网络(path

aggregation network, PAN) 的结构; Prediction 输出端由分类损失函数和回归损失函数组成, 用来评判检测效果。本文基于 YOLOv5 算法的人体跌倒检测方案的基本思路为: 首先分析 YOLOv5 算法的网络结构; 然后通过 YOLOv5 算法网络模型对标注的数据集进行训练; 最后对检测效果进行验证和分析, 如图 2 所示。

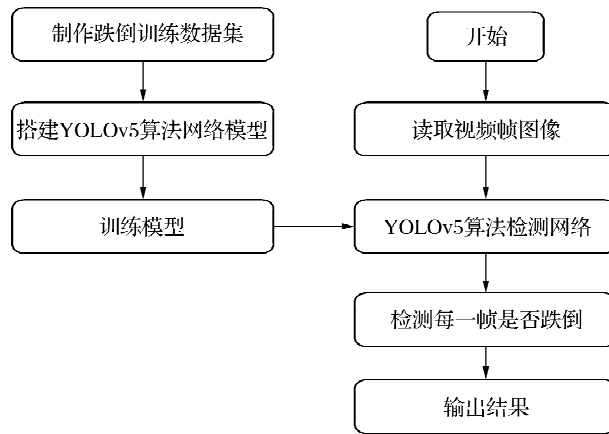


图 2 基于 YOLOv5 算法的人体跌倒检测方案

## 2 基于 YOLOv5 算法损失函数的训练和验证

### 2.1 训练集中 GloU 损失函数、目标检测损失函数、分类损失函数的训练结果分析

训练集中 GloU 损失函数、目标检测损失函数、分类损失函数的训练结果如图 3 所示。从图 3 可以看出: 在 0~10 迭代次数范围内, 三种类型的损失函数均值几乎呈垂直下降趋势; 在 10~200 迭代次数范围内下降幅度变缓, 表明训练结果理想。

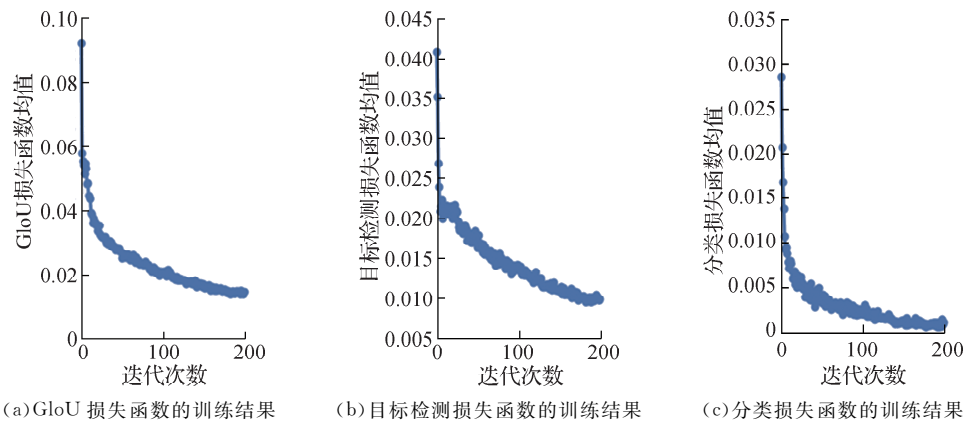


图 3 训练集中 GloU 损失函数、目标检测损失函数、分类损失函数的训练结果

这三种类型的损失函数均值经过 200 次迭代后, 其中 GloU 损失函数均值下降到 0.015~0.017, 表明最小外接矩形的贴合度和准确率均比较高; 目标检测损失函数均值下降到 0.01 左右, 表明目标检测准确率高; 分类损失函数均值下降到 0.001 左右, 表明分类准确率高。

### 2.2 验证集中 GloU 损失函数、目标检测损失函数、分类损失函数的验证结果分析

验证集中 GloU 损失函数、目标检测损失函数、分类损失函数的验证结果如图 4 所示。从图 4 可以看出: 在 0~10 迭代次数范围内, 三种类型的损失函数均值几乎呈垂直下降趋势; 在 10~200 迭代次数范围内, GloU 损失函数均值和分类损失函数均值继续缓慢下降, 而目标检测损失函数均值明显上升。

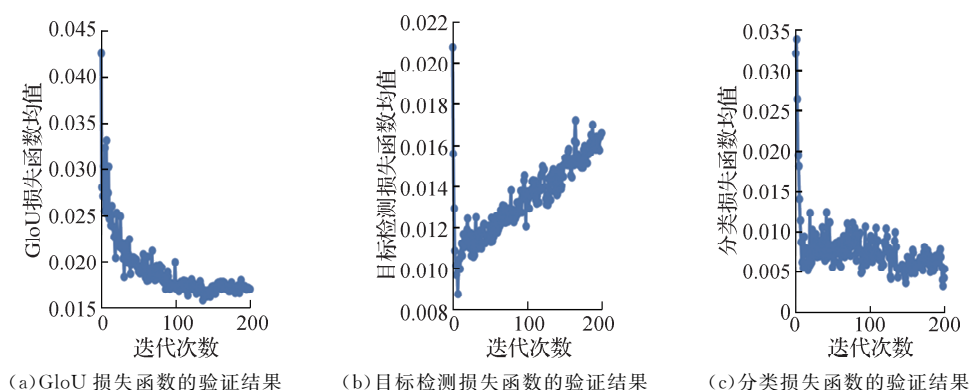


图 4 验证集中 GloU 损失函数、目标检测损失函数、分类损失函数的验证结果

这三种类型的损失函数均值经过 200 次迭代后, GloU 损失函数均值下降至 0.017 左右; 分类损失函数均值下降至 0.0026 左右, 表明最小外接矩形贴合度和分类准确率较高, 验证结果理想; 目标检测损失函数均值达到 0.0167 左右, 表明目标检测准确率较高。

### 3 参数的训练

#### 3.1 精确度、召回率的训练结果分析

训练结果主要观察精确度和召回率的波动情况, 如果波动不大则表示训练效果好。图 5 为精确度、召回率的训练结果。从图 5 可以看出: 在 0~10 迭代次数范围内, 精确度和召回率几乎呈垂直上升趋势; 在 10~200 迭代次数范围内, 精确度在 0.82 左右波动, 召回率在 0.78 左右波动, 波动范围不大, 表明训练效果好。波动是因为每次训练卷积结果不一样, 所以训练集中需要分出一部分到验证集中观察损失函数的变化情况, 当损失函数不再改变时, 表明此阶段是训练的最佳阶段。

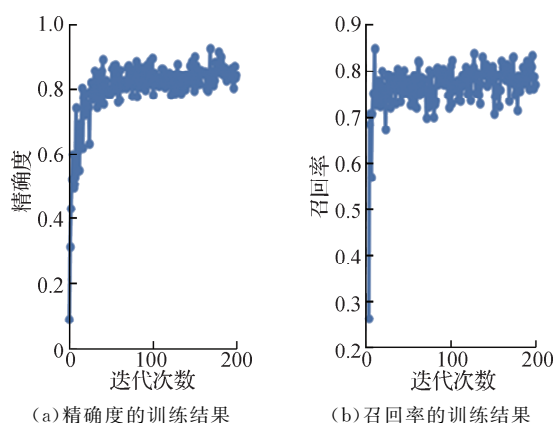


图 5 精确度、召回率的训练结果

#### 3.2 平均精确度均值(mAP)的训练结果分析

IoU 表示正负样本的阈值, 一般来说 IoU 阈值设置低, 样本的质量就会变差; IoU 阈值设置高, 样本的质量就会变好, 但是过高的 IoU 阈值会造成小尺度目标框的丢失。图 6 为不同 IoU 阈值下的 mAP 训练结果, 由图 6(a)可知: 当 IoU 阈值为 0.5 时, 在 0~10 迭代次数范围内, mAP 的值几乎呈垂直上升趋势; 在 10~200 迭代次数范围内, mAP 的值在 0.82 左右波动。

由图 6(b)可知: 不同 IoU 阈值(0.5、0.55、0.6、0.65、0.7、0.75、0.8、0.85、0.9、0.95)上 mAP 的平均值在 0~10 迭代次数范围内, 几乎呈垂直上升趋势; 在 10~200 迭代次数范围内, 上升幅度变缓, 经迭代 200 次后, mAP 的平均值在 0.68 左右波动。

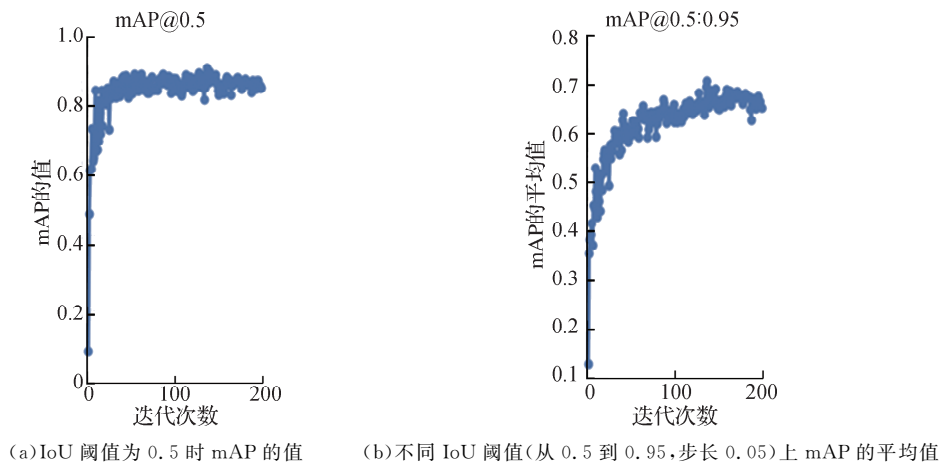


图 6 平均精确度均值(mAP)的训练结果

### 4 树莓派的实现

由于树莓派算力不够,可以通过加载 FastAPI 库编写代码来完成树莓派与电脑的连接。当程序运行时,树莓派将图片发送到电脑端,借助电脑的算力完成检测,电脑再将检测结果反馈给树莓派显示出来。图片中有多个或者单个目标人物站立、坐着或摔倒时,用本文算法检测都能被正确地识别出来并进行分类,具体的检测结果如图 7—图 9 所示。



图 7 目标人物站立时的检测结果



图 8 目标人物坐着时的检测结果

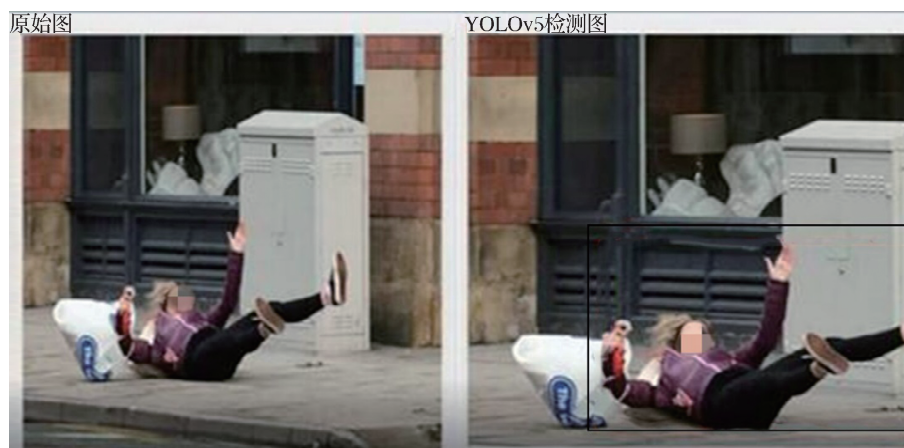


图 9 目标人物摔倒时的检测结果

## 5 实验结果分析

### 5.1 对站立、下蹲和跌倒动作识别精确度的比较分析

为了检验 YOLOv5 算法检测人体动作的真实精确度,共搜集了国内外 300 幅典型的图片来进行验证,检测出来跌倒动作的识别精确度能够达到 92%,站立动作的识别精确度能够达到 98%,下蹲动作由于处于跌倒和站立两动作之间,识别精确度仅有 82%。

### 5.2 不同光线环境下对目标人物动作识别置信度的比较分析

在光线充足和阴暗环境下的目标人物动作识别结果的比较分析分别如图 10 和图 11 所示。结果显示,光线充足时图片的置信度要高于昏暗环境时图片的置信度,表明光线充足与否会对目标人物动作检测造成一定的影响。

YOLOv5 算法只能检测出主要目标人物的动作,但并不能检测出边缘处一些小目标人物的动作,在训练时要针对一些小目标进行一定量的训练,如图 12 所示。



图 10 光线充足环境下的检测结果

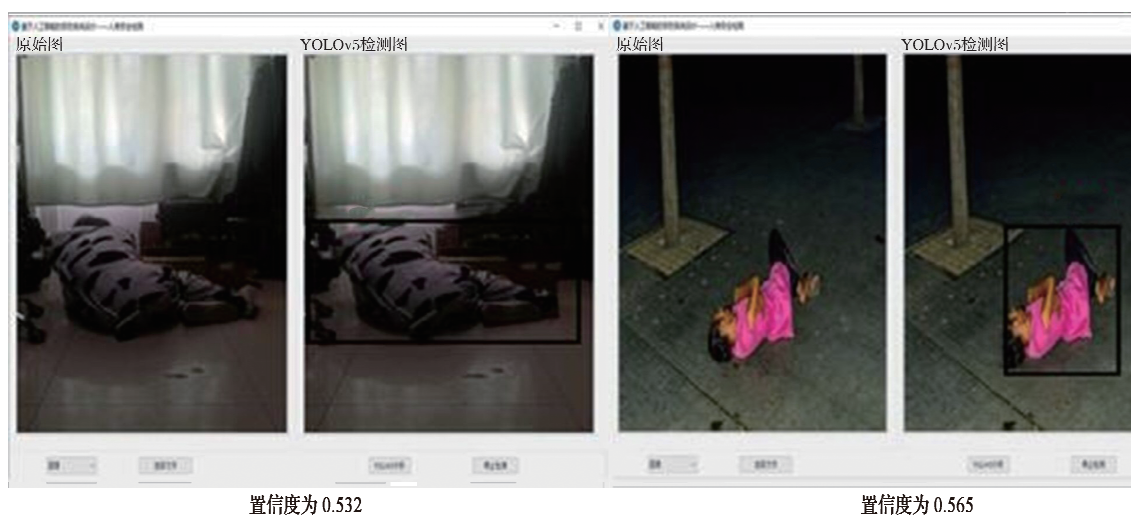


图 11 阴暗环境下的检测结果



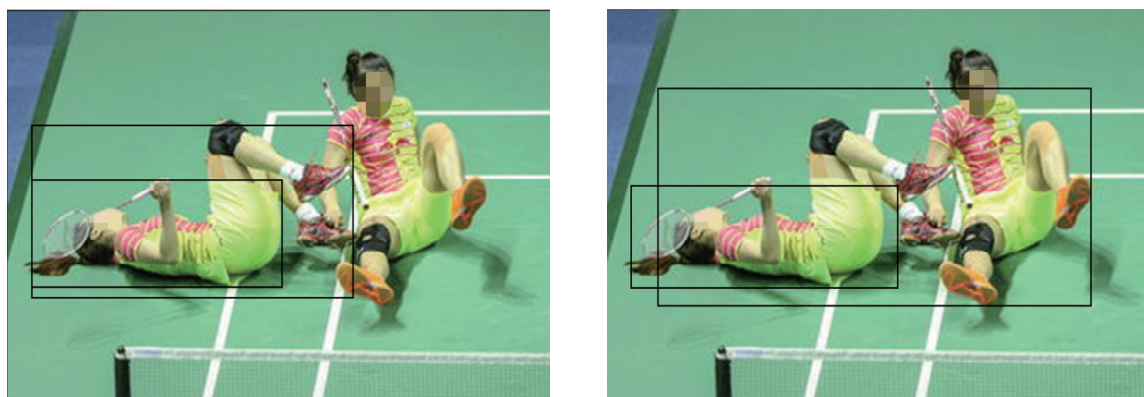
图 12 小目标的检测结果

### 5.3 算法对比分析

为了直观地了解 YOLOv5 算法检测目标人物动作的性能,选择 Faster R-CNN 算法进行对比。通过同一种数据集训练模型,观察两种算法的精确度以及性能指标,结果发现:1)YOLOv5 算法每帧检测只需要 0.016 s,而 Faster R-CNN 需要 0.785 s,因此 YOLOv5 算法比 Faster R-CNN 算法快得多(表 1),Faster R-CNN 算法不能满足实时检测需求;2)YOLOv5 算法 mAP 的值比 Faster R-CNN 算法高 36% (表 1);3)YOLOv5 算法每秒传输的帧数是 Faster R-CNN 算法的 2.25 倍;4)Faster R-CNN 算法对运行配置的要求要比 YOLOv5 算法高;5)YOLOv5 算法虽然不能识别出全部跌倒人员,但最小矩阵框选范围精确度很高,Faster R-CNN 算法虽然能够识别出全部跌倒人员,但最小矩阵框选范围不够精确,因此 YOLOv5 算法更适用于最小矩阵贴合度较高的目标检测(图 13)。

表 1 算法对比分析

算法	mAP 的值	每帧检测时间/s	每秒传输帧数(FPS)
YOLOv5	0.83	0.016	45
Faster R-CNN	0.61	0.785	20



(a)YOLOv5 算法的最小矩阵框选范围

(b)Faster R-CNN 算法的最小矩阵框选范围

图 13 YOLOv5 算法与 Faster R-CNN 算法的最小矩阵框选范围对比分析

## 6 结 语

本文基于 YOLOv5 算法设计了一种人体跌倒检测系统,能够检测出目标人物的跌倒、站立及下蹲动作。通过算法对比发现,YOLOv5 算法检测人体动作的速度和精确度高于 Faster R-CNN 算法,YOLOv5 算法更适用于最小矩阵贴合度较高的目标检测。虽然 YOLOv5 算法能够精确地检测出主要人物的动作,但不能识别出边缘上的一些小目标,因此在后续的研究中需要针对一些小目标进行一定量的训练。

### 参考文献:

- [1] 谷应雯,尚越.中国失能老人照护模式选择及其影响因素分析——基于非正式照护与正式照护的关系[J].卫生经济研究,2021,38(1):54-57
- [2] 李旭冬,叶茂,李涛.基于卷积神经网络的目标检测研究综述[J].计算机应用研究,2017,34(10):2881-2886
- [3] GUO Y L,WANG H Y,HU Q Y,et al. Deep learning for 3D point clouds;a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,43(12):4438-4364
- [4] QIU S,ANWAR S,BARNES N. Dense-resolution network for point cloud classification and segmentation[C]//Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa:IEEE,2021:3812-3821
- [5] CHEN C,FRAGONARA L Z,TSOURDOS A. GPointNet:graph attention based point neural network for exploiting local feature of point cloud[J]. Neurocomputing,2021,438:122-132
- [6] ISLAM K,DANG L M,LEE S J,et al. Image compression with recurrent neural network and generalized divisive normalization[C]//IEEE/CFV Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville:IEEE,2021:1875-1879
- [7] 徐毅,李蓓蓓,宋威.改进的深度置信网络分类算法研究[J].计算机科学与探索,2019,13(4):596-607
- [8] 吕璐,程虎,朱鸿泰,等.基于深度学习的目标检测研究与应用综述[J].电子与封装,2022,22(1):72-80
- [9] 刘颖,刘红燕,范九伦,等.基于深度学习的小目标检测研究与应用综述[J].电子学报,2020,48(3):590-601
- [10] 马健,阎卫东,刘国奇.基于 YOLOv5 的古建筑木结构裂缝检测方法[J].沈阳建筑大学学报(自然科学版),2021,37(5):928-934
- [11] 吴帅校. Faster R-CNN 目标检测网络算法压缩和 FPGA 实现[D].北京:北京交通大学,2021

(责任编辑:谭彩霞)