



计算机科学

COMPUTER SCIENCE

改进YOLOV7的跌倒人员检测

赵俊杰, 周晓静, 李佳欣

引用本文

赵俊杰, 周晓静, 李佳欣. [改进YOLOV7的跌倒人员检测](#)[J]. 计算机科学, 2024, 51(6A): 230800039-6.

ZHAO Junjie, ZHOU Xiaojing, LI Jiaying. [Improved YOLOV7 for Fall Detection](#)[J]. Computer Science, 2024, 51(6A): 230800039-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于改进YOLOv7的钢材缺陷检测](#)

Steel Defect Detection Based on Improved YOLOv7

计算机科学, 2024, 51(6A): 230800018-5. <https://doi.org/10.11896/jsjcx.230800018>

[基于SPD-Conv结构和NAM注意力机制的鱼群小目标检测](#)

Small Object Detection for Fish Based on SPD-Conv and NAM Attention Module

计算机科学, 2024, 51(6A): 230500176-7. <https://doi.org/10.11896/jsjcx.230500176>

[基于改进的YOLOv5s列车轨道线检测方法](#)

Railway Track Detection Method Based on Improved YOLOv5s

计算机科学, 2023, 50(11A): 230200101-6. <https://doi.org/10.11896/jsjcx.230200101>

[改进YOLOv5s的轻量化钢材表面缺陷检测模型](#)

Improved YOLOv5s Lightweight Steel Surface Defect Detection Model

计算机科学, 2023, 50(11A): 230900113-7. <https://doi.org/10.11896/jsjcx.230900113>

[基于深度学习的跌倒行为识别](#)

Fall Action Recognition Based on Deep Learning

计算机科学, 2019, 46(9): 106-112. <https://doi.org/10.11896/j.issn.1002-137X.2019.09.014>

改进 YOLOV7 的跌倒人员检测

赵俊杰 周晓静 李佳欣

河北工业职业技术大学宣钢分院 河北 张家口 075100

摘要 随着人口老龄化的到来,老年人在跌倒后能否及时发现并得到救治显得越来越重要。采用 YOLOV7 可以通过图像识别老年人跌倒,为提高原始 YOLOV7 的检测精度和速度,本研究对 YOLOV7 进行了一系列改进,并提出了一种新型 YOLOV7 结构——YOLOV7-CMJ。首先对收集到的图片进行处理,对部分图片进行了旋转、亮度等预处理,并对其进行标定以获取样本数据集;其次引入 CBAM 注意力机制,增强了通道注意力和空间注意力,从而提升模型的准确性;最后,将 YOLOV7 中原有的 PANet 特征融合改为 MJPANet,即多跳动征融合结构,并将之前的 Concat 采用加权的方式进行替换,从而得到改进 YOLOV7-CMJ 结构。通过与原始 YOLOV7 进行实验对比可知,改进后的算法精度提高了 7.4%、召回率提高了 7.1%、平均精度提高了 7.1%,证明了改进算法的有效性,更好地满足了摔倒检测要求。

关键词 跌倒检测;YOLOV7;跳动连接;CBAM;加权连接

中图分类号 TP391

Improved YOLOV7 for Fall Detection

ZHAO Junjie, ZHOU Xiaojing and LI Jiaying

Hebei Vocational University of Industry and Technology, Zhangjiakou, Hebei 075100, China

Abstract With the advent of the aging population, it is increasingly important for the elderly to be detected and treated in time after falling. In order to improve the detection accuracy and speed of the original YOLOV7, a series of improvements are made to YOLOV7, and a new YOLOV7 structure, namely YOLOV7-CMJ structure, is proposed. Firstly, the collected pictures are processed, and some pictures are preprocessed with rotation, brightness and other preprocessing, and calibrated to obtain sample datasets. Secondly, the CBAM attention mechanism is introduced to enhance channel attention and spatial attention, thereby improving the accuracy of the model. Finally, the original PANet feature fusion in YOLOV7 is changed to MJPANet, that is, multi-beating sign fusion structure, and the previous Concat is replaced by weighting, so as to improve the YOLOV7-CMJ structure. By comparing with the original YOLOV7, it can be seen that the accuracy of the improved algorithm is increased by 7.4%, the recall rate is increased by 7.1%, and the average accuracy is increased by 7.1%, which proves the effectiveness of the improved algorithm and better meets the requirements of fall detection.

Keywords Fall detection, YOLOV7, Jump connection, CBAM, Weighted joins

1 引言

研究表明,跌倒是全球范围内非故意伤害中最主要因素。老年人中,95%髌部骨折是由跌倒引起的^[1]。目前我国已经进入老龄化人口阶段,老年人口数量急剧增加,与此同时老年人跌倒后会有 17%的概率导致重大疾病^[2]。研究还表明,老年人跌倒后如果能及时得到救治可以降低 80%的死亡率,因此对跌倒后的老年人及时进行救治显得尤为重要^[3]。目前关于跌倒检测的方式主要有两种。一种是基于穿戴设备,主要有穿戴背心、穿戴头盔等。其主要是通过加速度计和陀螺仪测算速度、加速度,当速度加速度的值超过某一阈值时判定人员已经跌倒,穿戴检测设备伴随着传感器技术和通讯技术的不断迭代更新,其发展也较为迅速,并且具有设备种类繁多、在检测时不受环境限制等特点。其优点在于在收集用户运动数据的同时还可以保护个人隐私,并且可以在任意地方进行穿戴检测,使用不受场地限制,但是老年人有时会遗忘穿戴造成无法检测,而且穿戴设备成本高、续航时间短也是其必须面对的问题。另一种是通过计算机视觉的方法进行检测。首先

是通过将生活中采集到的跌倒图像进行标定和训练,将训练后的结果通过实时采集到的图像或视频处理来进行人体行为识别,从而判断是否发生了跌倒事件。该方法检测速度快精确度高,并且不受是否穿戴设备影响^[4]。

Xie 等利用 YOLOV7(You Only Look Once Version 7)对行人进行了跟踪,同时利用跟踪后的边界变化判断行人是否跌倒^[5]。Cheng 等利用 YOLOV3 锁定人体目标,结合 Alpha Pose 网络识别出骨骼关键点,并根据人体关节点的变化判定人员是否跌倒^[6]。Xu 等开发了基于 YOLOV5 的空巢老人跌倒检测系统,为提高跌倒检测的准确率和检测速度对模型进行了改进^[7]。本文选择使用较为先进的 YOLOV7 检测方法进行跌倒检测,同时基于 YOLOV7 的基本算法对其进行了改进,使其具有更高的检测精度和更快的检测速度。

2 添加 CBAM 注意力机制、MJPANet 网络和加权连接的改进 YOLOV7 结构

YOLOV7 算法是由 Chien-YAO 在 2022 年提出的一种较为新型的 YOLOV7 结构,YOLOV7^[8]和 YOLOV5^[9]相比

具有更高的检测精度和更快的检测速度。YOLOV7 和 YOLOV5 一样都是由输入端、主干网络、颈部网络和特征预测 4 部分组成,基本结构如图 1 所示。YOLOV7 的输入端是将图片像素调整为 640×640 大小的图片,之后将输入的图片传送到主干网络中,YOLOV7 相较于 YOLOV5 图片输入具有更高的分辨率,因此可以感受更多的图片细节,能够更好地分辨小目标。主干网络由 CBS 模块、ELAN 模块和 MPCConv 等模块组成,其中 CBS 模块是由 Conv 层、BN 层和 Silu^[10] 激活函数构成。YOLOV7 的 ELAN 模块不同于 YOLOV5 的 ELAN 模块,是一个更为高效的新型网络结构模块,该模块通过将不同路径的信息用不同的梯度路径向下传导并将各路信息连接在一起,从而使网络具有更强的适应性^[11]。MPCConv (混合卷积)由两条路径组成,一组是最大池化后用 CBS 模块进行处理,另一组是两次 CBS,最终将两组传出的信息特征进

行结合;颈部结构由 CSPNet(跨阶段局部结构)和 PANet(双向融合)结构组成,完成语义融合,其中包含 CBS 模块、UP-Sample 模块、SPPCSP(Spatial Pyramid Pooling Concat)模块、ELAN-H 模块和 REP 模块。UPSample 是向上采样的模块,它使用的上采样方式是最近邻插值;SPPCS 先通过 CBS 模块再通过最大池化,之后再通过 CBS 模块和之前的 CBS 模块相连接,进而可以获得不同的感受信息量,使得算法在不同的分辨率图像中都有较好的感受效果;ELAN-H 模块和 ELAN 模块基本机构类似,不同之处是 ELAN-H 输出数量不同。为了提高跌倒的识别精度和识别速度,以 YOLOV7 结构为基础,加入了 CBAM 注意力机制,这样使得模型可以获得更多的有效信息,同时采用的 MJPANet-多跳动加权特征融合使模型能够提取出更多的有用信息,以方便后面进行特征提取,最后将所有的连接替换成加权连接。

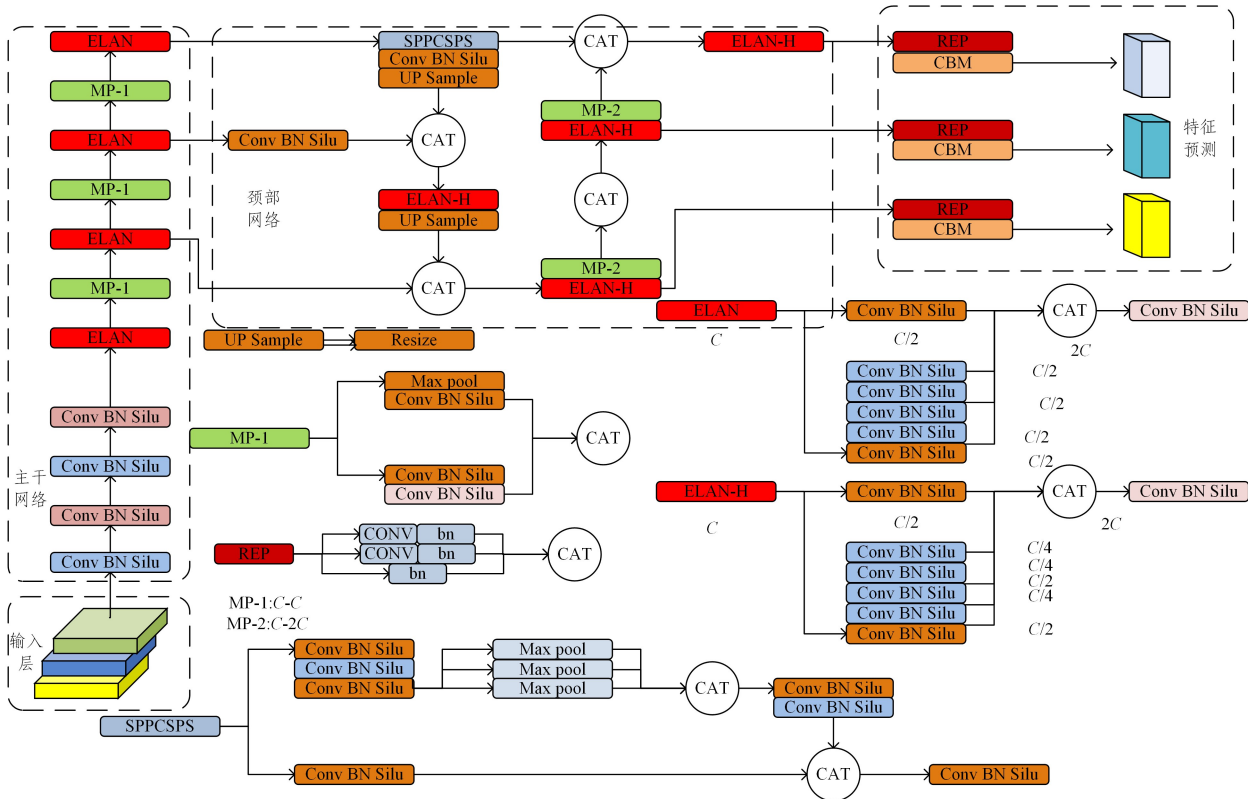


图 1 YOLOV7 算法模型

Fig. 1 YOLOV7 algorithm model

2.1 引入 CBAM 注意力机制

卷积块注意力模块 (Convolutional Block Attention Module, CBAM)^[12] 包含通道注意力模块 (Channel Attention Module, CAM)^[13] 和空间注意力模块 (Spatial Attention Mechanism, SAM)^[14] 两个子模块,其中 CAM 进行通道注意力计算,SAM 进行空间注意力计算。注意力机制是对图片中感兴趣和对识别有突出贡献的区域进行增强,以及抑制不感兴趣的区域,因此引入注意力模块可以对图像中的不同区域进行分割。由于注意力机制不改变传输通道中输入和输出信息的组成结构,只是改变其中的传输的特征,所有注意力机制可以做为即插即用的模块直接插入现有的 YOLO 结构中,并且能够节约参数和计算量^[15]。从图 2 中可以看出 CBAM 先输入一个特征,通道注意力接收到特征后对其进行全局最大池化和全局平均池化操作,将处理后的信息将其传递到下层;当空间注意力接收到信息进行平均池化、最大池化以及卷积

操作后再输出;之后通过 sigmoid 激活函数得出各个通道的权重因子,最终输出带有不同加权特征的特征图。

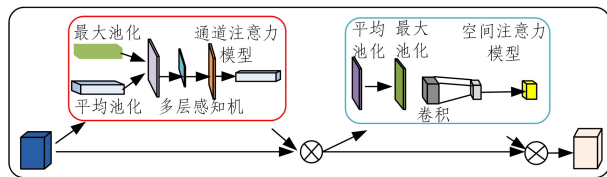


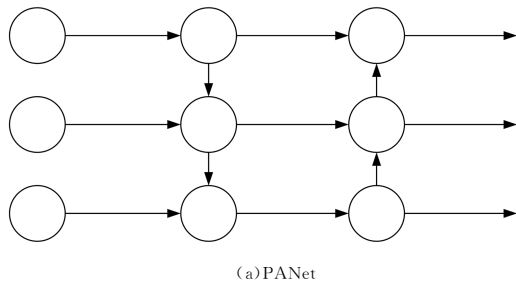
图 2 CBAM 整体结构图

Fig. 2 CBAM structure diagram

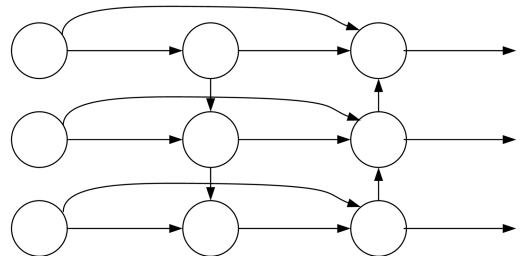
2.2 改进 PANet 为 MJPANet 加权多跳动网络结构

神经网络中的连接形式和组合结构对于信息的交互有着重要影响。为了获得更多的图片信息,同时增加信息的流动性和融合性,YOLOV7 和 YOLOV5 均采用 PANet(路径聚合网络)结构^[16],如图 3(a)所示,以此增加信息的交互和传递。

PANet 结构和原始结构的不同之处是在信息传递过程中加入了一条自顶向下和自底向上的信息传递路线,使得高层特征可以直接获得底层的更多信息^[17]。但是在这种结构中初始的信息特征在传递过程会部分丢失,原始信息的重要特征不能和最终输出建立直接关系。为了解决 PANet 连接中原始特征在传递过程中不能和最终输出连接的问题,对 PANet



(a) PANet



(b) MJPANet

图3 特征融合

Fig. 3 Feature fusion

2.3 加权连接

在特征传递过程中,YOLOV7 和 YOLOV5 都只是简单地用 Concat(连接)结构进行拼接,不能将每层结构进行高效的融合。对各层的重要性不能进行有效区分,使其在传递过程中加大了计算量,并且有效信息不能很好地传递给下层。为了解决 YOLOV7 特征融合只是简单的使用 Concat 对特征地图进行叠加连接,本文采用了简单而高效的加权特征融合机制,将各个特征进行加权叠加,加权之后更多的有效信息会得以保留,无效的信息在传递中会逐渐弱化,加权再叠加一方面可以对有效信息进行增益,另一方面减少无效信息可以减少每层的信息量。本文采用的是 Fast normalized fusion 加权方式,其计算公式为:

$$out = \sum_i \frac{w_i * I_i}{\epsilon + \sum_i w_i} \quad (1)$$

其中, w_i 权重代表数, I_i 为输入的特征图。为避免数值不稳定,引入 ϵ , out 为图 3 中每层对应的权重。

不同的特征和信息对于最终识别结果的贡献不同,通过将 CBAM 和加权连接引入 YOLOV7 中,可以提高对识别结果贡献大的特征和信息权重的传递,降低对识别结果贡献小的特征和信息权重的传递。改进的加权跳动可以将更多底层信息传递至第三层,加强了信息融合。综上所述,在以增加较小参数量为代价的前提下提高检测精度和速度。最终得到改进后的结构如图 4 所示。

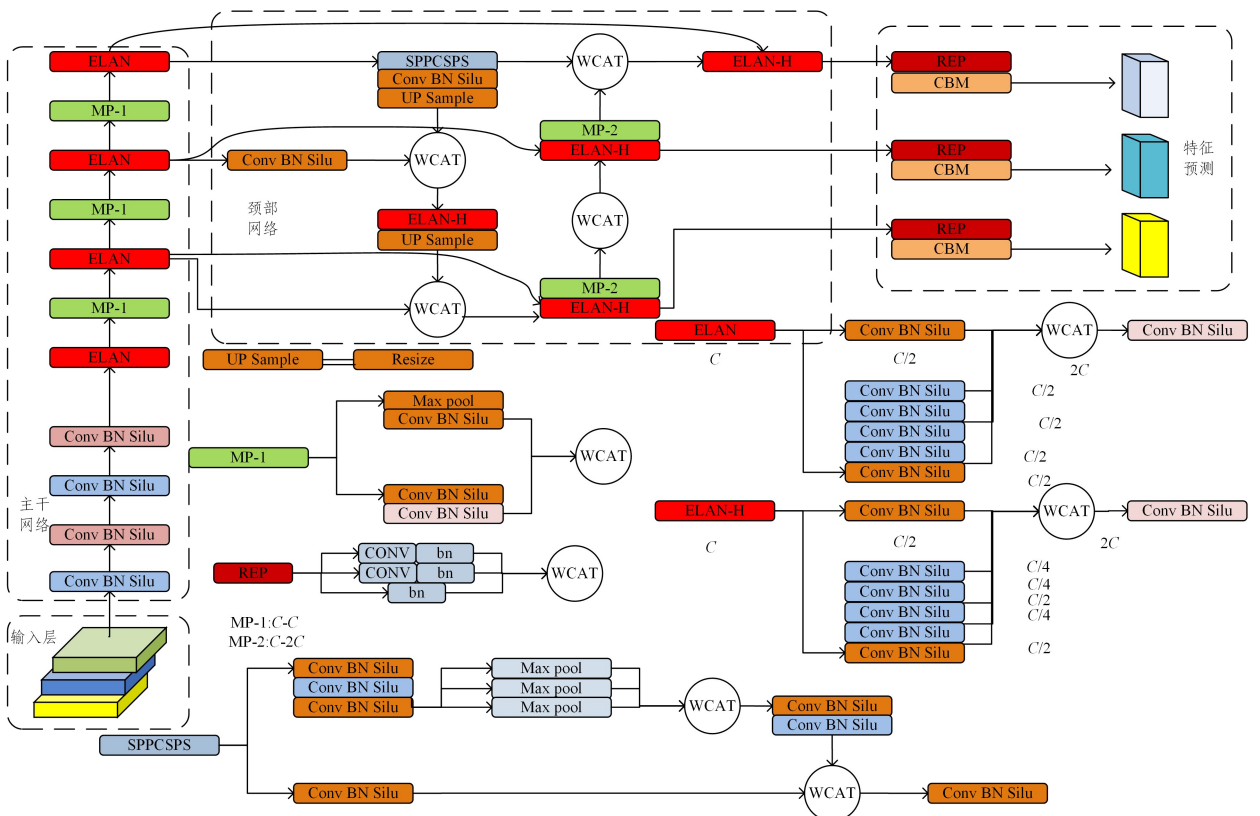


图4 yolov7CMJ 算法模型

Fig. 4 yolov7CMJ algorithm model

3 实验结果对比分析

3.1 实验环境

考虑到实际部署时为了降低成本,会对软硬件的性能进行限制,所以本文实验采用 Windows11 系统,AMD Ryzen 7 5800H 处理器,NVIDIA T600 Laptop GPU 显卡,Python3. 11, Pytorch1. 12,CUDA11. 3 软件。

3.2 实验所用数据及实验过程

本文采用的数据集一部分来自公开的跌倒数据集(UR-FD),另一部分通过网络进行图片收集,这样使数据集更接近真实生活下的各种跌倒场景,同时对部分图像进行了旋转、亮度变换以更接近现实拍照,最终得到 1000 张样本图像。使用 labimage 进行跌倒标定,将图片中的人物划分为倒下和非倒下两种状态,最后将图片训练集、测试集、检测集按照 8:1:1 的比例进行随机划分,如图 5 所示。

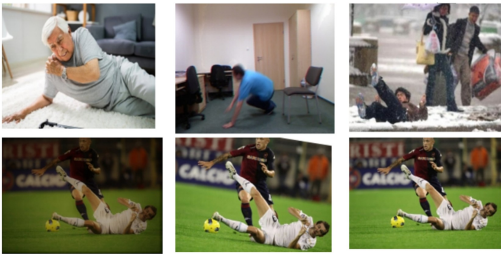


图 5 部分样本实例数据

Fig. 5 Partial sample instance data

在实验中设置初始学习率为 0.01,动量为 0.937,根据电

脑运行能力将批处理量设定为 3,图片分辨率为 640×640 ,将训练次数设定为 500 次,采用 SGD 优化器对参数进行迭代更新,通过 Warm-Up 方法预热学习率。

3.3 评价指标

在跌倒检测时,在考虑跌倒检测的准确度即能否正确的判断被检测者是否跌倒的同时,还需考虑检测速度,即及时做出跌倒判断同样很重要,所以对于新提出的模型主要采用以下评价指标对其进行评估,分别是精确率(P)、均值平均精度(mAP)、检测时间以及召回率(recall, R)和参数模型($Params$)^[18]。各检测指标的表达式分别如式(2)一式(4)所示:

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$mAP = \frac{1}{m} \sum AP(i) \quad (4)$$

其中, TP 代表的是正样本被模型识别的数量, FP 代表的是负样本被识别的数量, FN 代表正样本被错误识别的数量; AP 是以 P 为纵轴, R 为横轴构成 $P-R$ 曲线, P 为 $P-R$ 曲线下的面积, $P-R$ 曲线下的面积越大, AP 值越高,说明检测模型越准确, mAP 是所有类别 AP 的平均值, m 代表检测个数,如式(4)所示; $Params$ 为模型的参数量大小;通过计算检测时间可以直观反应出检测速度快慢。

3.4 消融实验

通过对原始 YOLOV7 加以改进可以提高其检测精度和速度。为了评估不同的改进模型以及不同模块组合后的结果差异(其结果如表 1 所列),用自建数据集进行了消融实验。

表 1 消融实验

Table 1 Ablation experiment

实验序号	CBAM	Mjpanet	Conunt 都加权	$P/\%$	$R/\%$	$Map@0.5\%$	$Params/M$	检测时间/s
1				0.925	0.7000	0.782	74.8	0.0621
2	✓			0.946	0.771	0.77	74.8	0.0273
3	✓		✓	0.928	0.743	0.791	74.8	0.0279
4			✓	0.926	0.756	0.787	74.8	0.0298
5		✓		0.980	0.714	0.814	74.8	0.0233
6		✓	✓	0.964	0.771	0.835	75.4	0.0229
7	✓	✓	✓	0.999	0.771	0.853	76.5	0.0289

通过消融实验可以看出不同模块的组成对检测结果会产生不同影响,例如:对比实验 1 和实验 2 可以看出,增加注意力机制后,模型量没有增加,精度增加了 2.1%,检测速度有所提升,但是召回率提升较大,同时平均精度又有所下降;对比实验 2 和实验 3 可以看出,增加权限后可以获得更多的有效信息,使平均精度有所提升,但误增加的更多错误信息致使召回率和精度下降;对比实验 1 和实验 4 可以看出,增加加权连接后,精度提升了 5.5%,召回率提升了 1.4%, mAP 增加 0.09%,同时检测速度也有所提升,虽然模型的连接数增多,但是模型量并没有增加,而且也取得了较好的训练结果;对比实验 1 和实验 5 可以看出,增加 Mjpanet 可以使召回率和平均精度都有所提高;在 Mjpanet 相同的情况下,增加全连接对比实验 5 和实验 6 的检测速度可以看出检测帧率有所下降,主要是加入 CBAM 模块后模型量有所增加,造成计算量有所增加;最后对比实验 7 和实验 1 可以看出,同时增加 CBAM 和 Mjpanet 以及连接都加权后,精度提升了 7.2%,召回率提升了 7.71%,平均提升了 7.71%,检测速度也有大幅度的提升。

由此可以证明,通过对 YOLOV7 进行一系列改进可以提高对应的检测准确率和检测速度。最终检测结果如图 6 所示。从图中可以看出该方法能够很好地检测出人体是否跌倒。



图 6 检测效果对比

Fig. 6 Comparison of detection effects

3.5 对比实验

目前主流的深度学习目标检测有 Two stage 算法和 One stage 检测方法。two stage 算法是两阶段检测,其主要包括 R-CNN^[19]算法、Fast R-CNN^[20]算法和 Faster R-CNN^[21]算法等,两阶段算法虽然检测精度高但检测速度较慢,无法实现实时性。one stage 是一阶段算法,其主要包括 SSD 算法^[22]和 YOLO 算法^[23-25]。一阶段检测可以直接定位到感兴趣点,所以检测速度较快,且经过近些年不断的优化改进,其检测精度与两阶段检测不相上下。为了验证改进后算法的有效性,本文将其与目前主流的 one stage 算法进行对比,即 SSD, YOLOV5, YOLOV7 同 YOLOV7CMJ 算法进行了对比。分析结果如表 2 所列。从表中可以看出 YOLOV7CMJ 具有最高的平均精度,其值分别较 SSD, YOLOV5L, YOLOV7 高出了 7.7%, 16%, 7.7%, 验证了算法的高效性。YOLOV7CMJ 的 map 精度相比原始 YOLOV7 提高了 7.7%, 同时训练时间较 YOLOV7 算法大幅减少。但由于相比原始 YOLOV7 算法加入了较多的连接,其模型大小比原始 YOLOV7 算法模型有所增加。在跌倒检测中存在着较多的躺卧图片,因为蹲下和跌倒也有部分相似,在考虑倒下姿势的同时还应考虑到沙发和床等背景,所以在进行人体跌倒检测时还需对背景有较为充分的理解才能够准确地判断出是否跌倒。本文中的跌倒平均精度已经达到 85.3%, 基本上达到检测要求。

表 2 对比实验

Table 2 Comparative experiment

实验模型	Map/%	Params	训练时间/h
SSD	0.787	91.1×10^6	3.120
YOLOV5	0.696	90.5×10^6	25.314
YOLOV7	0.782	74.8×10^6	30.178
YOLOV7CMJ	0.853	76.5×10^6	9.132

结束语 为提高对跌倒人员的检测精度和检测速度,本文对 YOLOV7 模型进行了改进,得到 YOLOV7CMJ 模型。基于目前较为先进的 YOLOV7 模型,加入 CBAM 注意力机制,使得特征能够更有效地将信息传递到下层,并且改进特征融合为多连接特征融合,这样使得原始信息和输出信息有更多的连接,最后将 Concat 结构都改为加权 Concat 结构,以方便信息的有效传递。最后通过实验对比发现,改进后的 YOLOV7CMJ 精度比未改进的模型提高了 7.4%、召回率提高了 7.1%、平均精度提高了 7.1%, 证明了其有效性,同时更好地满足摔倒检测要求。

参考文献

[1] United Nations Department of Economic and Social Affairs, Population Division. World population ageing 2020 highlights: living arrangements of older persons (ST/ESA/SER. A/451) [R/OL]. (2021-07-24) [2021-11-04]. www.un.org/development/desa/pd/.

[2] GE Y F, WANG L J, FENG W M, et al. The challenge and strategy selection of healthy aging in China[J]. Journal of Management World, 2020, 36: 86-95.

[3] HUANG M A, CHEN Y. The status quo and Advice of China's Ageing [J]. Economic Research Guide, 2018(10): 54-58, 66.

[4] ZHENG Y Z, ZHANG S. Review on classification of human fall detection technology [J]. Changjiang Information & Communications, 2021(1): 4. 15-18.

[5] XIE H, SHI H Q, QI Y X et al. Spationtemporal Fall Event Detection Algorithm Based on Attention Mechanism Subnetwork [J]. Computer and Modernization, 2022(3): 70-75, 81.

[6] CHENG S H, XIE W R, ZHANG D F, et al. Fall action recognition based on computer vision [J]. Acta Metrologica Sinica, 2022 (43): 107-113.

[7] XU W, LIAO Y K. Design and Research of Fall Detection System for Empty-nesters based on YOLOv5 [J]. Software Engineering, 2023, 26(2): 40-45.

[8] WANG C Y, BOCHKOVSKIY A, LIAO H. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[EB/OL]. [2022-07-06]. https://arxiv.org/pdf/2207.02696.pdf.

[9] JOCHER G. Yolov5[EB/OL]. (2020-08-10) [2020-08-10]. https://github.com/ultralytics/yolo-v5.

[10] ELFWING S, UCHIBE E, DOYA K. Sigmoid, weighted linear units for neural network function approximation in reinforcement learning[J]. Neural Networks, 2018, 107: 3-11.

[11] ZHANG X D, ZENG H, GUO S, et al. Efficient Long-Range Attention Network for Image Super-resolution[J]. arXiv: 2203.06697, 2022.

[12] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]// Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018: 3-19.

[13] FU J, ZHENG H, MEI T. Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2017: 4476-4484.

[14] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2018: 7132-7141.

[15] WANG F, WANG L, ZHANG R L, et al. Pedestrian Detection Algorithm Based on Fusion FPN and Faster R-CNN[J]. Journal of Data Acquisition and Processing, 2019, 34(3): 530-537.

[16] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 8759-8768.

[17] NIU Z Y. A review on the attention mechanism of deep learning [J]. Neurocomputing, 2021, 452: 48-62.

[18] PADILLA R, NETTO S L, DA SILVA E A B. A Survey on Performance Metrics for Object-Detection Algorithms[C]// 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). Niteroi, Brazil, 2020: 237-242.

[19] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vi-

- sion and Pattern Recognition. Columbus, OH, USA, 2014: 580-587.
- [20] GIRSHICK R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 1440-1448.
- [21] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [22] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// Proceedings of European Conference on Computer Vision. Washington D. C. , USA: IEEE Press, 2016: 21-37.
- [23] GE Z, LIU ST, WANG F, et al. YOLOX: Exceeding YOLO Series in 2021[J]. arXiv: 2107. 08430, 2021.
- [24] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. arXiv: 1804. 02767, 2018.
- [25] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [J]. arXiv: 2004. 10934, 2020.



ZHAO Junjie, born in 1989, master, lecturer. His main research interest is image recognition.