

# 基于深度学习的跌倒行为识别

马 露<sup>1</sup> 裴 伟<sup>2</sup> 朱永英<sup>3</sup> 王春立<sup>1</sup> 王鹏乾<sup>1</sup>

(大连海事大学信息科学技术学院 辽宁 大连 116026)<sup>1</sup>

(大连海事大学环境科学与工程学院 辽宁 大连 116026)<sup>2</sup>

(大连海洋大学海洋与土木工程学院 辽宁 大连 116026)<sup>3</sup>

**摘 要** 随着老龄人口的快速增长,跌倒检测成为医疗健康领域的一个关键问题。准确检测监控视频中的跌倒行为并及时反馈能有效减少老年人因跌倒造成的伤害甚至死亡。针对监控视频中的复杂场景及多种相似性人类行为干扰的情况,文中提出一种改进的 FSSD (Feature Fusion Single Shot Multibox Detector) 跌倒检测方法。首先,从不同的跌倒视频序列中抽取视频帧形成数据集;然后,将训练样本集输入到改进的 FSSD 网络中训练直至网络收敛;最后,根据最优化的网络模型测试视频中目标的类别并定位目标。实验结果表明,改进的 FSSD 算法可以有效检测每帧图像的跌倒或日常生活活动 (Activities of Daily Living, ADL) 事件并给出实时反馈,检测速度为 24 fps (GTX1050Ti),在保证检测精度的同时满足实时性要求。将改进方法与已有最新方法进行比较,结果表明:改进的 FSSD 算法的性能优于其他算法。视频中跌倒行为的检测进一步验证了基于深度学习的识别方法的可行性与高效性。

**关键词** 跌倒检测,卷积神经网络,FSSD 目标检测算法,深度学习,行为检测

中图分类号 TP391 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.09.014

## Fall Action Recognition Based on Deep Learning

MA Lu<sup>1</sup> PEI Wei<sup>2</sup> ZHU Yong-ying<sup>3</sup> WANG Chun-li<sup>1</sup> WANG Peng-qian<sup>1</sup>

(College of Information Science and Technology, Dalian Maritime University, Dalian, Liaoning 116026, China)<sup>1</sup>

(College of Environmental Science and Engineering, Dalian Maritime University, Dalian, Liaoning 116026, China)<sup>2</sup>

(Ocean and Civil Engineering Department, Dalian Ocean University, Dalian, Liaoning 116026, China)<sup>3</sup>

**Abstract** With the rapid growth of the aging population, fall detection has become a key issue in the medical and health field. Accurately detecting falling events in the monitoring video and giving feedback in real time can effectively reduce injuries even deaths caused by falls in the elderly. In view of the complex scenes in the monitoring video and multiple similar human behaviors, this paper proposed an improved FSSD (Feature Fusion Single Shot Multibox Detector) fall detection method. Firstly, a video frame forming dataset is extracted from different falling video sequences. Then, the training sample set is input into the improved convolutional neural network until the network converges. Finally, the target category and the location of the target in the video are tested according to the optimized network model. The experimental results show that the improved FSSD algorithm can effectively detect the falling or ADL activities of each frame of image and provide real-time feedback. The detection speed is 24 fps (GTX1050Ti), which can meet the real-time requirements while ensuring the detection accuracy. Comparing the improved method with the state-of-the-art fall detection methods, the performance of the improved FSSD is better than other algorithms. The detection of fall behavior in video further validates the feasibility and efficiency of the recognition method based on deep learning.

**Keywords** Fall detection, Convolutional neural network, FSSD target detection algorithm, Deep learning, Action detection

## 1 引言

跌倒是指突发、不自主的、非故意的体位变化,倒在地上

或者更低的平面上<sup>[1]</sup>。据世界卫生组织数据显示,每年约有 28%~35% 的 65 岁以上老人发生跌倒,70 岁以上老人跌倒的比例高达 32%~42%<sup>[2]</sup>。中国卫生部公布的《老年人跌倒

到稿日期:2018-07-09 返修日期:2018-09-21 本文受国家自然科学基金项目(61001158,61001158,61370070),辽宁省自然科学基金项目(2014025003),辽宁省教育厅科学研究一般项目(L2012270),大连市科技创新基金(2018J12GX043),辽宁省重点研发计划指导计划项目资助。

马 露(1993-),女,硕士生,主要研究方向为图像处理、异常行为识别,E-mail:malu930310@163.com;裴 伟(1977-),男,博士,副教授,主要研究方向为图像处理、数据挖掘,E-mail:peiwei2002@163.com(通信作者);朱永英(1978-),女,博士,副教授,主要研究方向为模式识别;王春立(1972-),女,教授,主要研究方向为模式识别与数据挖掘;王鹏乾(1997-),男,主要研究方向为电子信息工程、数据挖掘。

干预技术指南》指出,跌倒是我国伤害死亡的第四大因素,是 65 岁及以上老年人伤害死亡的第一大因素<sup>[1]</sup>。根据《中华人民共和国 2017 年国民经济和社会发展统计公报》数据显示,截至 2017 年底,我国 60 岁及 60 岁以上的老年人口占总人口的 17.3%<sup>[3]</sup>。老龄人口的快速增长,使得老年人的医疗健康成为社会关注的问题之一,而跌倒检测是医疗健康领域的一个关键问题。跌倒造成的伤害甚至死亡会给家人及社会造成大量的人力、财力损失。Noury 等的研究表明,老年人跌倒后如能得到及时救治,可以降低 80% 的死亡风险和 26% 的长期支援治疗风险<sup>[4]</sup>。因此,及时地检测出监控视频中的跌倒行为是十分重要的:一方面可以及时向监护人员发出警报寻求帮助,以减轻跌倒带来的伤害;另一方面可以节约社会公共医疗资源,减轻社会的公共医疗负担。

近十年来,研究者在跌倒检测算法方面已经开展了许多工作。根据使用设备和检测方法的不同,跌倒检测算法主要分为以下 3 个方面:基于可穿戴式传感器的方法、基于场景传感器的方法和基于计算机视觉的方法<sup>[5]</sup>。基于可穿戴式传感器的方法使用传感器自动检测跌倒,通过 WIFI、移动网、蓝牙等通信设备向医护人员发送求助信息<sup>[6-8]</sup>,因而具有很多优势。但是,老年人经常忘记佩戴传感器。基于场景传感器的方法利用安装在监控区域内的场景传感器采集压力、震动和声音等信息来判断跌倒是否发生。但是此类传感器对噪声信息比较敏感,容易产生误报<sup>[9-10]</sup>,并且布置成本较高,准确率较低。基于计算机视觉的跌倒检测方法通过从监控设备中被动获取人体运动信息,并对获取的视频或图像进行处理,来检测是否发生跌倒。基于计算机视觉的方法不需要用户随身佩戴任何设备,用户体验好,检测精度高。

在计算机视觉研究领域,卷积神经网络(CNN)等深度学习技术已经在图像分类方面广泛应用。CNN 能够直接从大量的标注数据中逐层抽取图像中的信息,提取图像的有效特征并对图像进行检测与分类。目前应用较广泛的多目标检测算法的网络模型有 Faster-RCNN<sup>[11]</sup>,YOLO<sup>[12]</sup>,SSD<sup>[13]</sup>等。Faster-RCNN 算法采用 RPN 网络提取 proposals。RPN 方法的本质是采用滑动窗口机制,RPN 在最后一个卷积层输出的特征图上进行滑窗,每个滑窗都产生 9 个 anchor,由于 anchor 机制和边框回归可以得到不同尺度、不同长宽比的候选区域,通过非极大值抑制(NMS)方法最终得到 300 个候选窗口进行检测。Faster-RCNN 需要对大量的候选窗口进行目标判定,再进行目标识别,实时性差<sup>[14]</sup>。YOLO 可以一次性预测多个候选框和目标类别,YOLO 将目标检测任务作为一个回归问题,真正意义上实现了端到端的目标检测,但是,它是牺牲精度来提高检测速度的。SSD 综合了 Faster-RCNN 与 YOLO 算法的优势,使用 anchor 机制代替 Faster-RCNN 的区域候选网络,同时采用了 YOLO 算法的回归思想。SSD 算法兼顾了检测精度与速度。因此,大量多目标检测算法工作基于 SSD 算法进行,如 DSSD<sup>[15]</sup>,RSSD<sup>[16]</sup>,FSSD<sup>[17]</sup>,DSOD<sup>[18]</sup>等。本文将在 FSSD 的基础上进行改进,使其适用于跌倒检测。

本文第 1 节介绍研究背景;第 2 节介绍相关工作;第 3 节介绍算法的改进;第 4 节给出实验结果与分析;最后总结全文。

## 2 相关工作

本节介绍基于计算机视觉的跌倒检测算法的相关工作。文献[19]详细地介绍了此方面的内容。根据使用算法的不同,现有跌倒检测方法主要分为以下两类,即基于传统机器学习的方法和基于深度学习的方法。

基于传统机器学习的方法通常包括前景目标分割、特征提取和分类训练 3 个部分。首先将目标与背景区分开来,提取人体姿态变化、质心变化速度、人体运动、头部运动轨迹等特征;然后用机器学习算法对提取的特征进行预测,区分跌倒事件和正常事件。传统机器学习方法主要受阴影、光照变化和遮挡的影响<sup>[20-21]</sup>,在复杂环境下的鲁棒性与可靠性较差,适用范围有限。

近十年来,深度学习<sup>[22]</sup>,特别是 CNN,在目标识别、图像分割、语音识别等领域取得了巨大成功。跌倒行为识别是医疗健康领域的一个重要分支,它的发展不仅能减少行为识别系统对人工的依赖,提高跌倒行为监测的效率,还可以为医疗报警系统、智能监控和智能家居等后续环节提供技术支持。因此,跌倒行为识别具有广阔的应用前景和巨大的经济价值,这促使研究人员将 CNN 应用于跌倒行为检测。Adhikari 等使用 Microsoft Kinect 获取 RGB-D 图像,使用背景减法对 RGB-D 图像进行矢量化的轮廓检测和图像表征来构建数据集,将构建的数据集作为 CNN 网络模型的输入数据<sup>[23]</sup>。经测试,该算法的精度提高到了 81%,但检测步骤繁琐,需要使用背景差分法区别环境与目标,然后判断目标是否跌倒。Huang 等<sup>[24]</sup>在 Two-Stream<sup>[25]</sup>的基础上提出使用 Three-Stream CNN,将 RGB 图像序列的空间信息、时间信息与人体运动流信息三者融合来识别跌倒行为,将识别精度提高到了 93.42%,但其所选用的数据集是非公共数据集。Min 等提出了一种基于深度学习的使用场景分析的跌倒行为检测方法,采用 Faster-RCNN 方法进行场景分析,在公共数据集上将精度提高到 95%<sup>[26]</sup>。Faster-RCNN 算法由于基于滑窗机制,速度仅为 5 fps,因此实时性较差。

该领域存在的问题大致分为以下几点:

(1)无适用于此类算法的公开数据集

公开数据集的图像标注格式与此类算法要求的格式不同,须对原数据集图像重新进行标注。现有跌倒检测算法中,通常使用较少量的图像进行评估,而在 CNN 训练中需要大量图像。在跌倒检测这一领域中,没有适用于此类算法的公开数据集直接作为训练数据集以供使用。

(2)工作环境

跌倒检测算法在复杂场景下的鲁棒性较差,存在遮挡、光照、阴影、多种相似性人类行为等干扰。

(3)跌倒算法的检测性能

由于对视频的分析具有时间和空间上的复杂性,而跌倒检测算法的理想情况就是尽可能地做到实时性和准确性,从而能够及时反馈给监护人员,使其迅速采取救援措施,因此,需要进一步提高跌倒事件的检测效率。

针对现有基于深度学习跌倒检测方法存在的问题,并结

合 FSSD 算法在行为识别上良好的表现,本文在 FSSD 算法的框架上进行改进,以提高跌倒检测算法在复杂环境下检测的鲁棒性与稳定性。

### 3 跌倒检测算法

#### 3.1 网络模型

SSD 采用多尺度的方法得到不同尺度的特征图,在不同尺度的特征图中选取不同尺度和宽高比的默认框(即单元格中一系列固定大小的虚线框形成的边界框),以增强默认框对检测目标形状的鲁棒性。如图 1 所示,检测图片中不同位置和大小目标对应特征图中不同尺度的默认框。

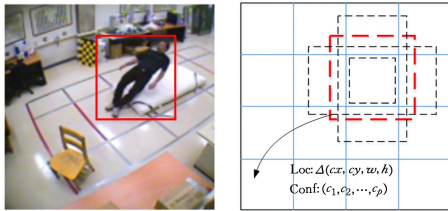


图 1 训练图片匹配

Fig. 1 Training image matching

在训练阶段,算法一开始会将这些默认框和真实框进行匹配。一个真实框可能对应多个默认框。在预测阶段,直接预测每个默认框的偏移值以及每个默认框中目标类别的相应

得分,使用 NMS 方法来移除得分较低的边框,并对类别相同且得分高的边框进行合并,最后通过 NMS 得到识别结果。SSD 模型的核心是预测目标类别,同时确定目标的位置。确定目标识别框是一个回归的过程,而确定识别框中的目标类别是一个分类的过程。因此 SSD 总的目标损失函数由位置损失与置信度损失的加权求和得到:

$$L(x, c, l, g) = \frac{1}{N} (L_{\text{conf}}(x, c)) + \alpha L_{\text{loc}}(x, l, g) \quad (1)$$

其中,  $N$  是与真实框相匹配的默认框个数,  $L_{\text{conf}}$  为置信度损失值,  $L_{\text{loc}}$  为位置损失值,  $x$  为输入图像,  $c$  为目标置信度,  $l$  为预测值,  $g$  为真实框,  $\alpha$  为权重项。

FSSD 网络模型基于 SSD 网络模型进行改进。FSSD 网络模型采用基础网络和辅助网络来提取图像特征。FSSD 网络使用 VGG16 作为基础网络初步提取图片特征。为获取更多小尺度下的信息, FSSD 在基础网络之后又依次增加 conv6\_2 和 conv7\_2 用于生成目标检测的特征图。将 conv4\_3, fc7 和 conv7\_2 输出的特征图调整为相同尺寸后进行特征融合, 利用融合得到的特征重构一组金字塔特征图(见图 2 中 Pyramid feature map)来进行预测。相对于传统 SSD 网络模型, FSSD 网络模型具有捕捉局部细节特征和全局语义特征的能力, 能通过增加上下文信息来提高跌倒检测的准确率。

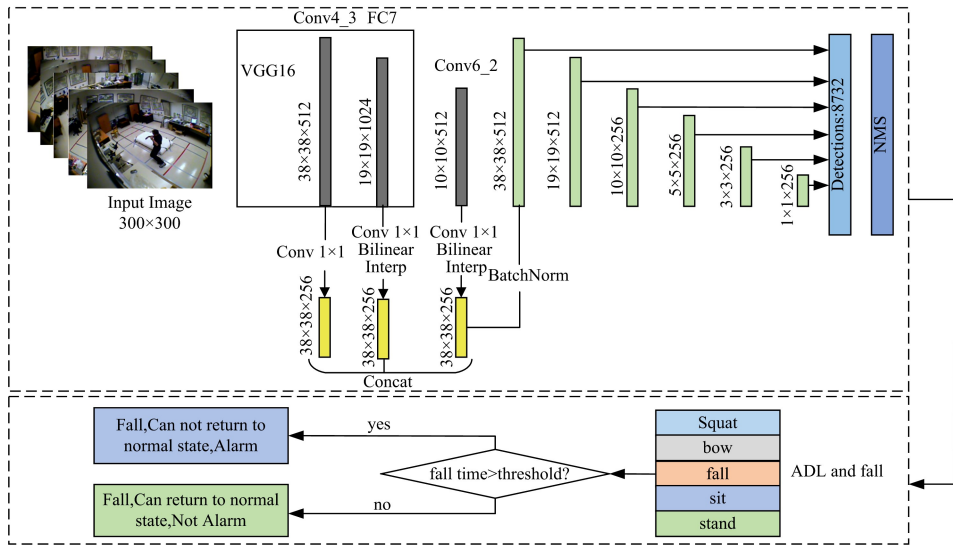


图 2 改进的 FSSD 算法结构

Fig. 2 Structure of improved FSSD algorithm

#### 3.2 特征融合模块

本文在 FSSD 算法框架的基础上,做出了如下改进:

(1) 移除部分卷积层以提高检测速度。增加的计算成本归因于滤波器的数量、网络深度、较小的步长及其组合<sup>[27]</sup>。由于时间成本受到限制,须将滤波器的数量、大小、深度、步长这些因素进行权衡<sup>[27]</sup>。本文把网络结构中的 conv7 层移除,并将 conv6\_2 的步长  $stride$  由 1 设置为 2,以保证特征图尺寸不小于  $10 \times 10$ 。经本文验证,尺寸小于  $10 \times 10$  的特征图几乎没有可用的信息进行融合。在神经网络结构中,更多的网络层意味着更多的参数,更多的参数意味着更多的计算,更大的内存消耗以及更低的速度。因此本研究选择更少的卷积层

以提高跌倒检测的速度,尽可能在不牺牲检测精度的同时提高检测速度。

(2) 选择 conv4\_3, fc7 和 conv6\_2 输出的特征图进行特征融合。SSD 基础网络中最小尺度的目标检测主要是由 conv4\_3 层产生的特征图决定的, conv4\_3 与其他层相比具有不同的特征尺度。多组对比实验结果表明: conv4\_3 与 conv6\_2 连接具有更好的性能。因此通过对特征图进行降采样或上采样,采用双线性插值将 fc7 和 conv6\_2 层的特征图尺寸统一为  $38 \times 38$  (conv4\_3 特征图的大小也被作为图像尺寸的临界值)。

(3) 提高模型检测精度后,通过跌倒时长来判断是否需要发出警报。如果在设定阈值内目标能够自行恢复正常状态,

则不需要发出警报;如果在阈值内目标未能自行恢复正常状态,则需要及时发出警报。

本文做出的修改如下:首先考虑到目标检测的实时性要求,在尽可能不损失检测精度的同时提高检测速度,本文选择更少的卷积层以提高检测速度;其次对 conv4\_3,fc7 和 conv6\_2 输出的特征图归一化后进行融合,以充分提取图像特征;最后通过跌倒时长判断是否能自行恢复正常状态,从而有效减少资源的浪费。

### 4 实验结果与分析

本实验对不同数据集中的跌倒目标进行检测,并对检测精度与检测速度进行验证。

实验的硬件配置为 Intel Core i5-7300HQ 处理器, NVIDIA Geforce GTX 1050Ti 显卡, 2.5 GHz CPU。软件环境为 Ubuntu14.04,cuda8.0,Opencv2.4.10,caffe 框架。

#### 4.1 数据集

本研究在两个公开的跌倒数据集上对图像进行重新标注、训练并验证。

Multiple cameras fall dataset(Multiple)包含用 8 台网络摄像头记录的 24 个场景。其中前 22 个场景包含跌倒和正常的日常活动,最后 2 个场景只包含正常的日常活动。每帧的大小为 720×480,帧率为 30 fps。本文随机选用 6150 张帧图像制作数据集,根据图片数量将训练集、验证集、测试集的比例设为 8:1:1。UR Fall Detection Dataset(URFD)是不同视角的户外摄像机拍摄的人类活动序列的集合,包含 30 个跌倒和 40 个日常生活视频片段,此数据集采用两个 Microsoft Kinect 相机拍摄。本文仅使用 30 个跌倒序列的 RGB 图像和 28 个正常活动序列。由于 URFD 数据集的数量大且标注格式与本文方法要求格式不同,因此本文随机选用 5510 张帧图像制作数据集,训练集、验证集、测试集的比例设为 8:1:1。

由于数据集包含多种复杂场景,并且存在很多干扰因素,如遮挡、光照变化、相机运动、多种人类行为等,因此检测上述数据集中的跌倒行为有很大挑战。

#### 4.2 评价标准

在目标识别领域,准确率和召回率是两个基本评价指标。根据准确率与召回率,可绘制准确率-召回率曲线(P-R 曲线),如图 3 所示。首先计算每一个类别的平均精度(Average Precision, AP),即 P-R 曲线下的面积;然后取多个类别平均精度的平均值作为平均精度均值(mean Average Precision, mAP)。本文将 mAP 作为跌倒识别模型的评价标准。

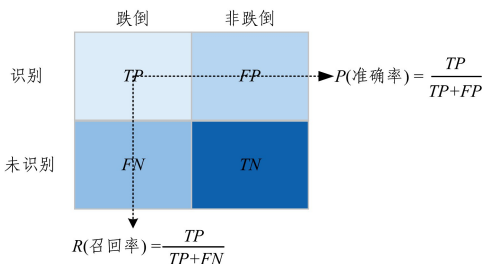


图 3 准确率与召回率示意图  
Fig. 3 Precision and recall rate

本文使用每秒检测帧数(FPS)评价目标检测速度,以 25 fps 作为衡量实时性的临界值。

图 3 中,TP(true positives)为本来是跌倒且分类成跌倒的次数;FP(false positives)为本来是非跌倒但分类成跌倒的次数;FN(false negatives)为本来是跌倒但分类成非跌倒的次数;TN(true negatives)为本来是非跌倒且分类成非跌倒的次数。

#### 4.3 实验设置

本文利用跌倒图像训练改进的 FSSD(Improved FSSD)网络模型来检测视频中的跌倒事件。算法框架在 Image 数据集上进行模型预训练,然后使用跌倒数据集微调预训练模型参数文件。

本文使用了随机梯度下降法,设定初始学习率为 0.0005,动量(momentum)为 0.9,权重衰减(weight\_decay)为 0.0005,批大小(batch\_size)为 8,学习率下降策略(lr\_policy)为多步策略(multistep)。正负样本由其归属类别的真实边框(ground truth)与预测边框的交并比 IOU(Intersection-Over-Union)决定,本文将 IOU 阈值设置为 0.55,以区分正负样本。根据权重文件调整参数,以使模型最优化。

具体实施步骤如下:首先对公共数据集中的图片进行筛选,并进行图片标注,按照 VOC2007 数据集的格式制作数据集;其次将训练集输入到 Improved FSSD 网络训练,根据权重文件调整网络结构与参数,直至网络收敛;最后使用最优化网络模型进行跌倒检测。

#### 4.4 实验结果

本文通过两组实验来验证改进的跌倒检测模型的可行性和高效性。

##### 4.4.1 Multiple cameras fall dataset

图 4 展示了 Improved FSSD 方法在 Multiple cameras fall dataset 上的部分检测结果,结果显示本文方法可从多角度、远距离(距离摄像头的距离)检测跌倒行为。

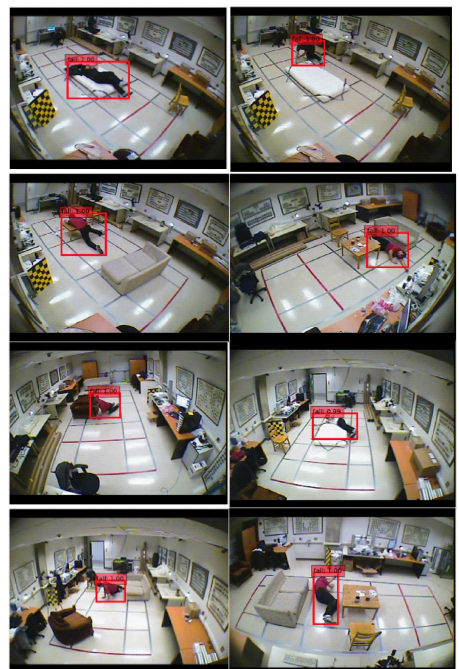


图 4 部分检测结果  
Fig. 4 Partial test results

图5展示了迭代次数与  $mAP$  之间的关系。由图5可知, Improved FSSD 算法的  $mAP$  值相比 FSSD 算法提高了 1.7%, 相比 SSD 算法提高了 4.8%。

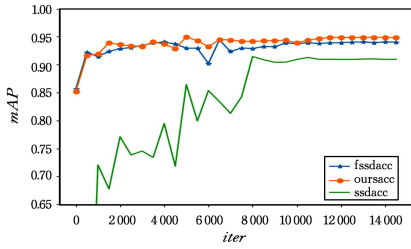


图5 迭代次数与  $mAP$  值的关系

Fig. 5 Relationships between number of iterations and  $mAP$  values

准确率和召回率是相互影响的, 理想条件下是二者都高, 但是一般条件下准确率与召回率成反比。如图6所示, 在跌倒检测方面, 当召回率增加到接近1时, 准确率有所下降。

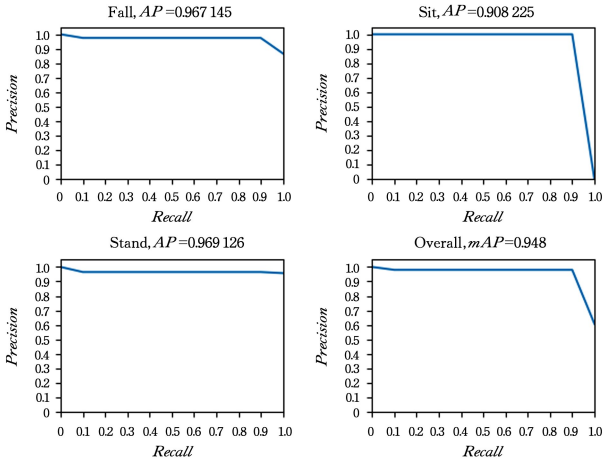


图6 P-R 曲线图

Fig. 6 P-R curve graph

表1统计了不同方法在不同检测目标上的  $mAP$  值。在跌倒这一类别中, Improved FSSD 的  $mAP$  值达到了 96.7%, 具有更好的检测能力。表2为本文方法与其他最新方法的比较, 在跌倒这一类别中, 本文方法的性能略优于其他算法, 这说明本文方法更加稳定、鲁棒性更高。

表1 不同算法的检测结果(Multiple)

Table 1 Test results of different algorithm (Multiple)

方法	$mAP$ /%	跌倒/%	坐下/%	站立/%	速度/Ips
SSD	90	90.1	87.2	92.7	27~28
FSSD	93.1	95.2	88.7	95.4	20~21
Improved FSSD	94.8	96.7	90.8	96.9	23~24

表2 本文方法与其他最新方法的性能比较

Table 2 Comparison of performance of proposed method and other state-of-the-art methods

方法	检测准确率
Min W et al. [26]	95.5
Rougier C et al. [28]	95.4
Mirmahboub B et al. [29]	95.2
Improved FSSD	96.7

(单位: %)

在视频检测时, 检测速度达到了 24 fps(GTX1050Ti), 与原始 SSD 方法相比, 虽然速度较慢, 但精度较高。由于我国采用 PAL 制式, 而 PAL 制式的帧速率为 25 fps, 因此本文方法能够基本满足实时性要求。

4.4.2 UR Fall Detection Dataset

图7展示了 Improved FSSD 算法在 UR Fall Detection Dataset 上的检测结果, 不同颜色框定了不同的行为类别。结果表明, 本文算法能正确识别目标类别并确定目标位置。图8展示了 Improved FSSD 算法检测视频的效果。图8(a)显示在第3帧时目标为正常活动状态, 图8(b)显示目标在第105帧时发生跌倒, 当跌倒这一动作的持续时间超过设定阈值时, 发出警报。在视频检测时, 检测速度达到了 24 fps(GTX1050Ti)。

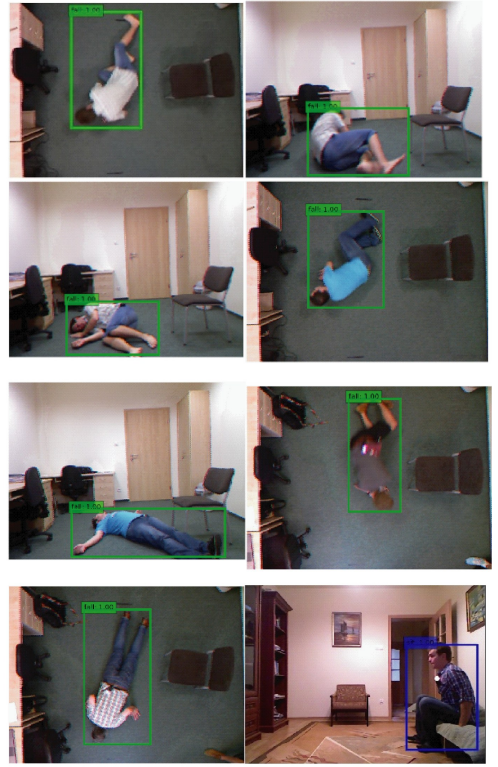


图7 部分检测结果

Fig. 7 Partial test results

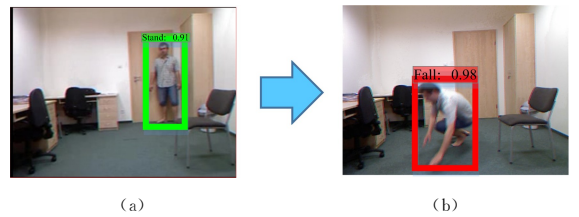


图8 视频检测结果

Fig. 8 Video test results

图9为本文算法与 FSSD 和 SSD 算法的精度值对比。由图9可知, Improved FSSD 与 FSSD 都有较高的检测精度, 在识别跌倒这一类别中, Improved FSSD 算法的  $mAP$  值比 FSSD 算法的  $mAP$  值提高了 2.23%, 比 SSD 算法的  $mAP$  值提高了 5.4%, 这说明 Improved FSSD 算法具有更高的准确

率。在站立和坐下这两类中, Improved FSSD 算法的检测精度均有明显提升,这说明本文方法在 UR Fall Detection Dataset 上具有更好的表现。本文方法的检测精度优于 FSSD 算法,这进一步说明了本文算法的稳定性与鲁棒性。

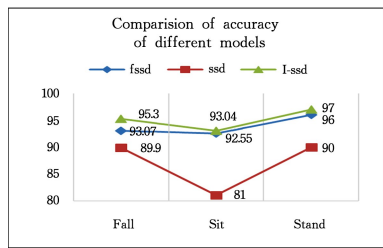


图 9 不同模型的目标检测结果

Fig. 9 Object detection results of different models

图 10 为 UR Fall Detection Dataset 的 P-R 曲线,由图可知,在跌倒这一类别中,本文算法在确保精确度较高的情况下,有较高的召回率。

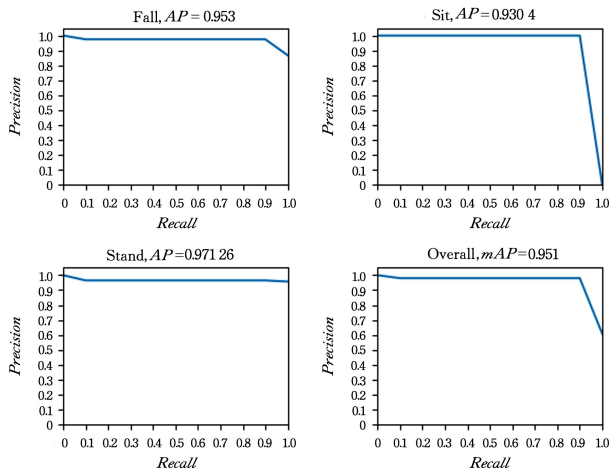


图 10 UR Fall 数据集上的 P-R 曲线

Fig. 10 P-R curve on UR Fall dataset

**结束语** 本文首先论述了跌倒检测的重要性。然后结合深度学习在跌倒检测方面的应用,分析了现有跌倒检测方法存在的不足。通过对几种广泛使用的多目标网络模型进行对比发现, FSSD 网络模型在性能上具有较大优势, FSSD 方法加入了特征融合模块,能通过增加上下文信息来提高目标检测的精度。

在此基础上,本文提出了一种 Improved FSSD 方法来检测跌倒行为。实验结果表明, Improved FSSD 方法可以有效检测每帧图像的 ADL 或跌倒行为并给出实时反馈。本文算法在 Multiple cameras fall dataset 上的检测精度为 94.8%,在 UR Fall Detection Dataset 上的检测精度为 95.1%,检测性能均优于其他方法,进一步证明了该方法具有更高的鲁棒性与稳定性。

本文改进的检测算法还有很多优化空间,在以后的研究中将继续改进模型,增强其共享能力,改善其实时性,同时结合 Atrous 滤波器、更换基础网络等多方因素来提高算法的检测性能。

## 参 考 文 献

- [1] PENG Y A. A Technical Guide to Falling Interventions for the Elderly[J]. Journal of Practical Rural Doctors, 2012, 19(8): 1-13. (in Chinese)  
彭宇案, 老年人跌倒干预技术指南[J]. 中国实用乡村医生杂志, 2012, 19(8): 1-13.
- [2] SANNINO G, FALCO I DE, PIETRO G DE. A supervised approach to automatically extract a set of rules to support fall detection in an mHealth system[J]. Applied Soft Computing, 2015, 34(C): 205-216.
- [3] National Bureau of Statistics of the People's Republic of China. Statistical Communique of the 2017 National Economic and Social Development of the People's Republic of China[J]. China Statistics, 2018(3): 7-20. (in Chinese)  
中华人民共和国国家统计局. 中华人民共和国 2017 年国民经济和社会发展统计公报[J]. 中国统计, 2018(3): 7-20.
- [4] NOURY N, FLEURY A, RUMEAU P, et al. Fall Detection—principles and Methods[C]// 29th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society. Lyon, France, 2007: 1663-1666.
- [5] MUBASHIR M, SHAO L, SEED L. A survey on fall detection: Principles and approaches[J]. Neurocomputing, 2013, 100: 144-152.
- [6] RATHI N, KAKANI M, EL-SHARKAWY M, et al. Wearable low power pre-fall detection system with IoT and bluetooth capabilities[C]// IEEE National Aerospace and Electronics Conference (NAECON). Dayton, OH, 2017: 241-244.
- [7] HOSSAIN F, ALI M L, ISLAM M Z, et al. A direction-sensitive fall detection system using single 3D accelerometer and learning classifier[C]// International Conference on Medical Engineering, Health Informatics and Technology (MediTec). Dhaka, 2016: 1-6.
- [8] SCHWICKERT L, BECKER C, LINDEMANN U, et al. Fall detection with body-worn sensors?: a systematic review [J]. Zeitschrift Für Gerontologie Und Geriatrie, 2013, 46(8): 706-719.
- [9] YAZAR A, ÇETIN A E. Ambient assisted smart home design using vibration and PIR sensors[C]// 21st Signal Processing and Communications Applications Conference (SIU). Haspolat, 2013: 1-4.
- [10] ARSHAD A, KHAN S, ALAM A H M Z, et al. A capacitive proximity sensing scheme for human motion detection[C]// IEEE International Instrumentation and Measurement Technology Conference (I2MTC). Turin, 2017: 1-5.
- [11] REN S, GIRSHICK R, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look

- Once, Unified, Real-Time Object Detection[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, 2016:779-788.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[J]. arXiv:1512.02325v2, 2015.
- [14] ZHOU X Y, WANG K, LI L Y. Review of object detection based on deep learning[J]. *Electronic Measurement Technology*, 2017(11):89-93. (in Chinese)
- 周晓彦, 王珂, 李凌燕. 基于深度学习的目标检测算法综述[J]. *电子测量技术*, 2017(11):89-93.
- [15] FU C, LIU W, RANGA A, et al. DSSD: Deconvolutional single shot detector[J]. arXiv:1701.06659, 2016.
- [16] JEONG J, PARK H, KWAK N. Enhancement of ssd by concatenating feature maps for object detection[J]. arXiv:1705.09587, 2017.
- [17] LI Z, ZHOU F. FSSD: Feature fusion single shot multibox detector[J]. arXiv preprint arXiv:1712.00960, 2017.
- [18] SHEN Z, LIU Z, LI J, et al. DSOD: Learning Deeply Supervised Object Detectors from Scratch[C]// Proceedings of the IEEE International Conference on Computer Vision. 2017:1937-1945.
- [19] ZHANG Z, CONLY C, ATHITSOS V. A survey on vision-based fall detection[C]// 8th ACM International Conference on pervasive Technologies Related to Assistive Environments. 2015:1-7.
- [20] MITERAN J, DUBOIS J, ATRI M. Optimized spatio-temporal descriptors for real-time fall detection: comparison of support vector machine and Adaboost-based classification[J]. *Journal of Electronic Imaging*, 2013, 22(4):41106.
- [21] YUN Y, INNOCENTI C, NERO G, et al. Fall detection in RGB-D videos for elderly care[C]// 17th International Conference on E-health Networking. Boston, MA, 2015:422-427.
- [22] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553):436.
- [23] ADHIKARI K, BOUCHACHIA H, NAIT-CHARIF H. Activity recognition for indoor fall detection using convolutional neural network[C]// Fifteenth IAPR International Conference on Machine Vision Applications. Nagoya, 2017:81-84.
- [24] HUANG C D, WANG C Y, WANG J. Human action recognition system for elderly and children care using three stream ConvNet [C]// International Conference on Orange Technologies. Hong Kong, 2016:5-9.
- [25] SIMONYAN K, ZISSERMAN A. Two-stream convolutional networks for action recognition in videos[J]. *Advances in Neural Information Processing Systems*, 2014, 1(4):568-576.
- [26] MIN W, 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, PP(99):1.
- [27] HE K, SUN J. Convolutional neural networks at constrained time cost[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015:5353-5360.
- [28] ROUGIER C, MEUNIER J, ST-ARNAUD A, et al. Robust Video Surveillance for Fall Detection Based on Human Shape Deformation[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011, 21(5):611-622.
- [29] MIRMAHBOUB B, SAMAVI S, KARIMI N, et al. Automatic monocular system for human fall detection based on variations in silhouette area[J]. *IEEE Transactions on Biomedical Engineering*, 2013, 60(2):427-436.