

面向社区医疗的跌倒检测算法

赵普¹, 武一^{1,2}

1. 河北工业大学电子信息工程学院, 天津 300401; 2. 河北工业大学电子与通信工程国家级实验教学示范中心, 天津 300401

【摘要】为解决社区中的独居老人发生跌倒时不能及时得到救治造成二次伤害的问题,提出一种面向社区医疗的跌倒检测算法。算法具有2D卷积和3D卷积两个分支,同时进行空间特征和时序特征的提取。在3D分支引入密集连接,增强时序特征提取能力;在2D分支重新设计网络的残差块,增强空间特征提取能力。在两个分支融合部分引入非局部注意力机制,增强特征融合能力。在算法的最后融合场景信息,通过SIOU损失函数和联合损失函数监督算法检测跌倒行为。在扩充后的公开URFD数据集上进行对比实验,本文算法检测准确率为98.3%,表明所提算法对于跌倒行为有较好的表现性和鲁棒性。

【关键词】跌倒检测;非局部注意力机制;损失函数;社区医疗

【中图分类号】R318;TP391.4

【文献标志码】A

【文章编号】1005-202X(2023)12-1486-08

Fall detection algorithm for community healthcare

ZHAO Pu¹, WU Yi^{1,2}

1. School of Electronics Information Engineering, Hebei University of Technology, Tianjin 300401, China; 2. National Experiment Teaching Demonstration Center for Electronic and Communication Engineering, Hebei University of Technology, Tianjin 300401, China

Abstract: A fall detection algorithm for community healthcare is proposed to avoid the secondary injury caused by untimely treatment when the elder living alone falls in the community. The algorithm has two branches, namely 2D convolution and 3D convolution, which allow it can extract spatial and temporal features simultaneously. The dense connections added in the 3D branch enhance the ability to extract temporal features; the residual blocks in the 2D branch are redesigned to improve the ability of spatial feature extraction; and a non-local attention mechanism is introduced to the branch fusion for better feature fusion. The algorithm also takes scene information into consideration, and it is supervised by SIOU loss function and the combined loss function to realize fall detection. The experiment on the expanded public URFD dataset reveals that the proposed method has a detection accuracy of 98.3%, which verifies its performance and robustness for fall detection.

Keywords: fall detection; non-local attention mechanism; loss function; community healthcare

前言

随着我国人口老龄化趋势日益加剧,独居老人的健康安全成为人们关注的焦点问题。老人跌倒后及时救治可有效降低老人的死亡风险^[1]。因此,及时发现独居老人跌倒成为研究的重点。目前,国内外跌倒检测的方法主要有3种类型,分别为基于可穿戴

式设备、环境传感器、计算机视觉。基于可穿戴式设备的方法需佩戴相应设备,极易造成不适,忘记佩戴时,也不能及时进行监测^[2-3];基于环境传感器的方法虽然无需穿戴,但是易受到周围环境的干扰,造成误报,准确率低,并且布置价格昂贵^[4-5]。随着计算机视觉不断的发展,深度学习被应用到跌倒检测中,只需架设一台摄像机,将采集的图像经过深度学习网络进行特征提取,使用相应的损失函数使得网络收敛,从而判断是否跌倒。

基于深度学习的方法不会影响人体活动,具有实时性好、成本低、检测准确度高的特点。曹建荣等^[6]融合运动特征和深度学习进行跌倒检测。首先通过改进的YOLOv3^[7]算法检测人体目标,然后得到运动特征,最后将其与YOLOv3算法融合。马露等^[8]

【收稿日期】2023-07-07

【基金项目】国家自然科学基金(51977059);河北省自然科学基金(E2020202042)

【作者简介】赵普,硕士,研究方向:计算机视觉,E-mail: zp_dling@163.com

【通信作者】武一,博士,教授,研究方向:智能系统控制,E-mail: wuyi-hbgdx@163.com

基于SSD(Single Shot MultiBox Detector)^[9]提出一种改进的SSD跌倒检测网络,对单帧跌倒图像进行检测。文献[6, 8]仅对跌倒过程的不同状态进行检测,割裂了跌倒行为的连续性,时序特征不足,易与类似跌倒行为混淆,产生误检。因此,有研究将3D卷积应用于跌倒检测^[10]。李秀智等^[11]提出一种轻量级的3D卷积神经网络跌倒检测算法,并引入显著性特征提取,加强背景纹理与人物行为之间关联性以提高真实场景的跌倒检测准确度,但该算法不能检测出跌倒者位置。有研究通过判断姿态行为进行跌倒检测。Li等^[12]融合2D姿态和3D姿态,经过长短时记忆网络进行时序特征提取,判断跌倒,但该算法模型较复杂,实时性较低。马敬奇等^[13]通过轻量化的Alphapose^[14]姿态估计算法计算人体关节之间的线速度和夹角变化来判断跌倒。孙颖等^[15]使用OpenPose^[16]提取人体骨架后计算人体躯干以及腿部与地面夹角,通过夹角判断是否发生跌倒。由于人体姿态复杂多变,仅对关节点进行物理判断,易产生错误检测。

本研究面向社区医疗养老提出一种融合时序特征和空间特征的老人跌倒检测算法,同时使用2D卷积与3D卷积进行空间特征和时间特征的提取。时间

特征的提取克服了割裂跌倒行为连续性的缺点,提高检测准确度;空间特征的提取克服使用3D卷积进行跌倒检测时,不能对跌倒人员定位的缺点。在特征融合部分引入非局部注意力机制^[17],非局部注意力机制通过全局建模增强特征融合能力。在算法的最后与场景信息融合以降低误检率,并使用SIOU损失函数^[18]作为预测框损失函数对跌倒行为进行检测。

1 算法设计

眼睛观察事物时,首先接收的是一帧画面,并将每一帧画面相互连接,形成连续帧。眼睛所接收的每一帧画面同时会与前面帧的画面联系起来。受此启发,本研究网络由2D卷积分支和3D卷积分支构成主干网络,2D卷积输入的是单帧RGB图像,3D卷积输入的是采样频率为1的16帧图像。在特征融合部分将2D卷积提取的单帧图像的空间特征与3D卷积提取的16帧图像的时序特征进行融合,特征融合部分由扩张卷积和非局部注意力机制共同构成,通过空间特征获取场景信息,场景信息包括人体宽高比、人体质心高度和人体倾斜角,并将场景信息与算法融合得到最终的检测结果。算法结构图见图1。

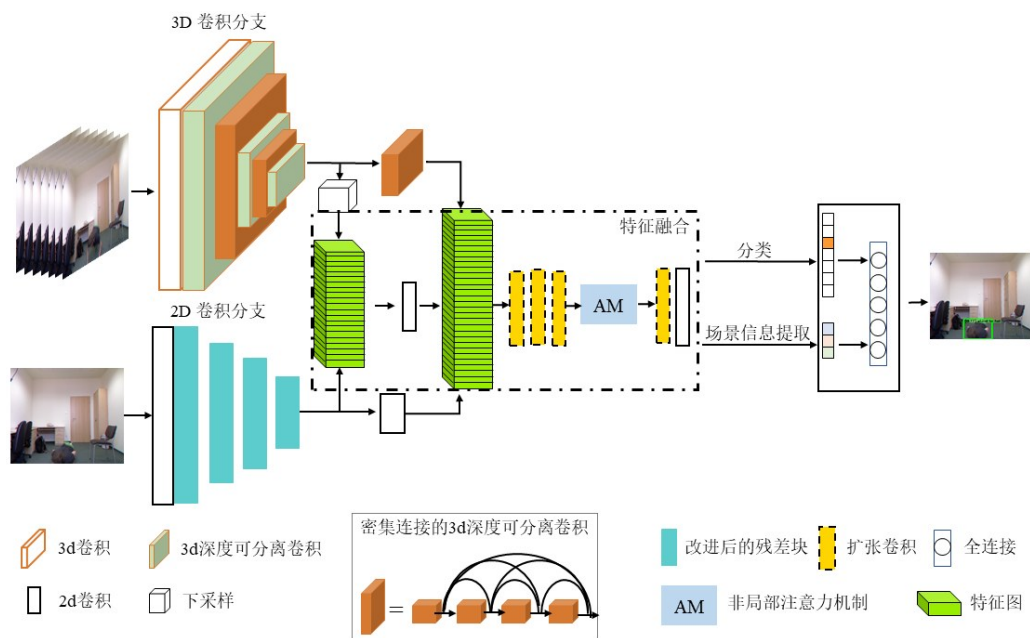


图1 算法结构图
Figure 1 Algorithm structure

1.1 3D卷积分支

由于3D卷积参数量较大,为便于部署,3D卷积分支由3D深度可分离卷积构成,并进行密集连

接^[19],使得下一层卷积层可以接收之前的所有层的特征作为额外的输入,实现特征的重用,在参数和计算成本更少的情况下提升网络的时序特征提取能

力。3D卷积分支的具体结构见表1,其中3d_conv为标准的3D卷积,conv3d_dw为3D深度可分离卷积,dense_block为密集连接的3D深度可分离卷积。

表1 3D卷积分支结构表
Table 1 3D convolution branch structure

输入大小	卷积层名字	卷积操作	步长	输出大小
224×224×3	3d_conv	3×3×3 conv	1, 2, 2	112
112×112×16	conv3d_dw	3×3×3 conv_dw/1×1×1 conv	1, 1, 1	112
112×112×32	conv3d_dw	3×3×3 conv_dw/1×1×1 conv	2, 2, 2	56
56×56×32	conv3d_dw	18*{3×3×3 conv_dw}/3*{1×1×1 conv}	1, 1, 1	56
56×56×48	conv3d_dw	3×3×3 conv_dw/1×1×1 conv	2, 2, 2	28
28×28×64	dense_block1	24*{3×3×3 conv_dw}/6*{1×1×1 conv}	1, 1, 1	28
14×14×96	conv3d_dw	3×3×3 conv_dw/1×1×1 conv	2, 2, 2	14
14×14×128	dense_block2	18*{3×3×3 conv_dw}/6*{1×1×1 conv}	1, 1, 1	14
7×7×192	conv3d_dw	3×3×3 conv_dw/1×1×1 conv	2, 2, 2	7
7×7×240	dense_block3	9*{3×3×3 conv_dw}/3*{1×1×1 conv}	1, 1, 1	7

1.2 2D卷积分支

2D卷积分支使用YOLOv4-tiny^[20]网络的Darknet特征提取部分时易产生漏检(图2a),为使网络能获取充分的空间特征,降低漏检率,重新设计YOLOv4-tiny的残差块部分,不再进行卷积块重复堆叠,并命名为mix_resblock。如图2b所示,其中CBL表示为卷积、归一化和LeakyReLU激活函数3步操作。首先对输入的特征图使用3×3卷积进行残差连接,并对特征图进行shuffle^[21]操作,使得原始的特征图之间进行信息交互。再使用5×5、3×3两个不同大小的卷积核进行特征提取。然后将不同卷积核大小的特征图进行软阈值化后与原始特征图进行拼接。软阈值化的过程是将绝对值低于某阈值的特征置为0,首先对特征的绝对值进行全局平均池化,再进行全连接、归一化、全连接操作。最后使用Sigmoid函数进行阈值限制。改进后的残差卷积块使用不同大小的卷积核使得训练更有效,学习更多样的特征。同时使用软阈值化保留有效特征图。

1.3 特征融合

首先将14×14的特征图进行融合,经过下采样后与2D卷积分支的7×7特征图和3D卷积分支的7×7特征图进行堆叠。将时序特征五维张量转换为四维张量,3层扩张卷积对特征图进行编码,相比于普通卷积,扩张卷积可以在不增加参数量的情况下提供更大的感受野,以便各个特征层间进行更好的融合。为避免栅格效应,本研究采用不同尺寸的扩张卷积,3层扩张卷积的尺寸分别为(1, 2, 5)。然后融入非局

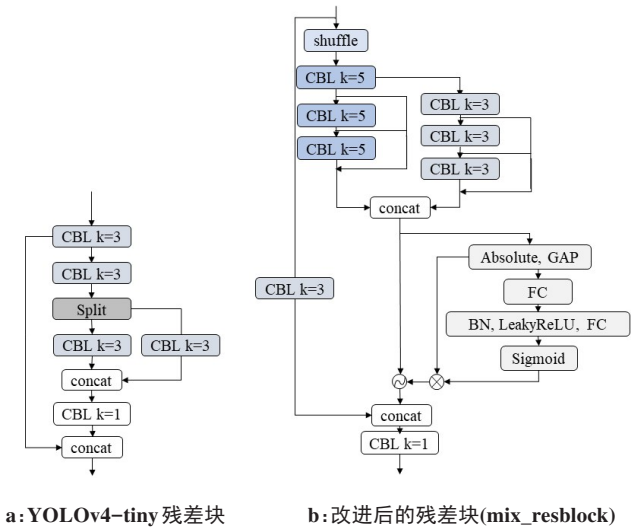


图2 YOLOv4-tiny resblock和mix_resblock对比图
Figure 2 Comparison of YOLOv4-tiny resblock and mix_resblock

部注意力机制,通过计算任意两个位置之间的空间信息并捕捉远程依赖性,提高不同分支之间特征图的融合能力。其公式如下所示:

$$y_i = \frac{1}{C(x)} \sum_{j \in V} f(x_i, x_j) g(x_j) = \frac{1}{C(x)} \sum_{j \in V} e^{\theta(x_i)^T \phi(x_j)} W_g x_j \quad (1)$$

其中,i为计算输出位置的索引,j是枚举所有可能位置的索引;x为输入特征图矩阵,y为对应的输出特征图矩阵;函数f计算i和j之间的空间相似性。这里函数选为高斯嵌入函数,即f(x_i,x_j)=e^{θ(x_i)^Tφ(x_j)}。其中θ(x_i)=W_θx_i和θ(x_j)=W_φx_j是两个嵌入特征矩阵,分别由输入x与1×1卷积得到。C(x)为归一化因

子,此时 $C(x) = \sum_{j \in V} f(x_i, x_j)$ 。函数 g 计算在位置 j 的输入特征图的表示,为线性嵌入形式,即 $g(x_j) = W_g x_j$ 。 W_θ 、 W_ϕ 、 W_g 均为可学习的 1×1 卷积。

1.4 损失函数

网络的损失函数分别为回归损失、分类损失、置信度损失和联合损失。

对于回归损失,使用 SIOU 损失函数对预测框进行监督, SIOU 相比 CIOU 在距离、形状和 IoU 代价函数的基础上增加预测框与真实框的角度偏移,使得模型更快速准确地收敛。损失函数的参数的含义见图 3。其中, B 表示预测框, B^{gt} 为真实框, σ 为真实框与预测框的中心距离, α 为真实框与预测框的偏移角,则 $\beta = \pi/2 - \alpha$ 。 C_w 、 C_h 是以 σ 为对角线的矩形的宽高, c_w 、 c_h 为最大外接矩形的宽高, w 、 h 为对应框的宽高。

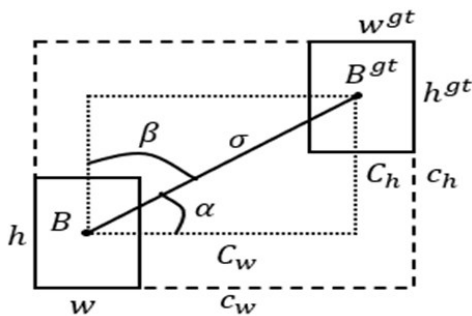


图3 SIOU参数图

Figure 3 SIOU parameters

SIOU 公式如下所示:

$$L_{SIOU} = 1 - \text{IoU} + \frac{\sum_{t=x,y} (1 - e^{-(2-\Lambda)\rho t}) + \sum_{t=w,h} (1 - e^{-\theta t})}{2} \quad (2)$$

(b_{cx}, b_{cy}) 为对应框中心点的坐标。 θ 对于每个数据集都是唯一的,决定了形状代价函数重要程度,实验表明 θ 的设置范围为 2~6,本研究中, θ 设置为 4。 Λ 为增加的角度代价函数,公式如下:

$$x = \sin(\alpha) = \frac{c_h}{\sigma} \quad (3)$$

$$\Lambda = 1 - 2 * \sin^2\left(\arcsin(x) - \frac{\pi}{4}\right) \quad (4)$$

新增加的角度代价函数使得距离代价函数能更快收敛,当 α 逐渐趋近于 0 时,距离代价函数的监督作用逐渐减弱,当 α 趋近于 $\pi/4$ 时,距离代价函数的监督作用达到最强,使得整个损失函数难以收敛。因此,应尽可能使 α 趋于 0。

对于分类损失,使用 focal loss 作为分类损失函数见式(5):

$$L_{\text{focal}}(p_t) = -\alpha_t (1 - P_t)^\gamma \log(P_t) \quad (5)$$

α 解决类别不平衡的问题,当某个类别过高时,可以减小 α 值,反之亦然。 γ 使得模型更关注错分的样本。 P_t 是模型的预测结果, P_t 越大则说明模型的训练的结果较好,反则说明预测结果较差。实验时,设置 $\alpha = 0.25$, $\gamma = 2$ 。

置信度损失是判断有无目标对象,计算预测框与真实框之间的损失,计算公式如下所示:

$$L_{\text{conf}} = \lambda_{\text{noobj}} \sum_{i=0}^{L_h * L_n} \sum_{j=0}^{L_n} I_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{obj}} \sum_{i=0}^{L_h * L_n} \sum_{j=0}^{L_n} I_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \quad (6)$$

当检测到有目标时, λ_{noobj} 为 0,反之则为 1。

回归损失和分类损失相互独立地进行网络优化,彼此之间缺少相关性,使得分类任务和回归任务预测不一致。分类任务只会选出类别得分最高的样本,不会考虑回归任务的 IoU 是否最高,回归任务也是如此,这就会导致网络选出得分最高的 IoU,而分类得分较低的样本被抛弃,干扰后续检测结果。为此,设计联合损失函数,使回归分支和分类分支进行联合优化,公式如下所示:

$$L = L_{\text{conf}} + (1 + \phi_f) L_{\text{SIOU}} + (1 + \phi_c) L_{\text{focal}} \quad (7)$$

其中, $\phi_f = e^{-L_{\text{focal}}}$ 和 $\phi_c = e^{-L_{\text{SIOU}}}$ 为两个联合因子。使得分类损失和回归损失相互监督,网络具有更好的表现力。

1.5 场景信息融合

由 2D 卷积分支检测出前景目标人物,并在其矩形框内取得前景目标人物的外接椭圆,外接椭圆相较于矩形框能有效地描述前景目标人物的姿态特征,因为肢体的伸展会给矩形拟合带来较大的噪声,不能准确地描述场景信息。外接椭圆的参数为椭圆中心点 $O(x, y)$,椭圆的长半轴 a 和短半轴 b ,定义长半轴 a 与短半轴 b 的比值为人体宽高比 β_1 ,即 $\beta_1 = a/b$;椭圆中心点到水平线 l 的距离为人体质心高度 h ;椭圆长轴与 y 轴的夹角为人体倾斜角 θ_1 。场景信息参数图见图 4。

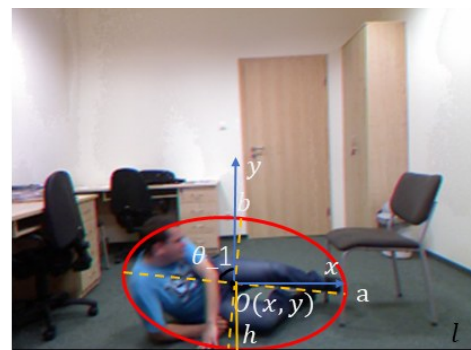


图4 场景信息参数图

Figure 4 Scene parameters

为客观地评价以上场景信息的特征变化情况,分别选取站立跌倒、坐下和行走状态的视频进行分析,场景信息变化曲线见图5。当站立跌倒时,人体宽高比值逐渐增大,人体质心高度逐渐减小,人体躯

干倾斜角逐渐增大;当行走和坐下时,人体宽高比和人体躯干倾斜角变化微小。跌倒时的场景信息变化有利于跌倒检测,且有效区别于行走和坐下。

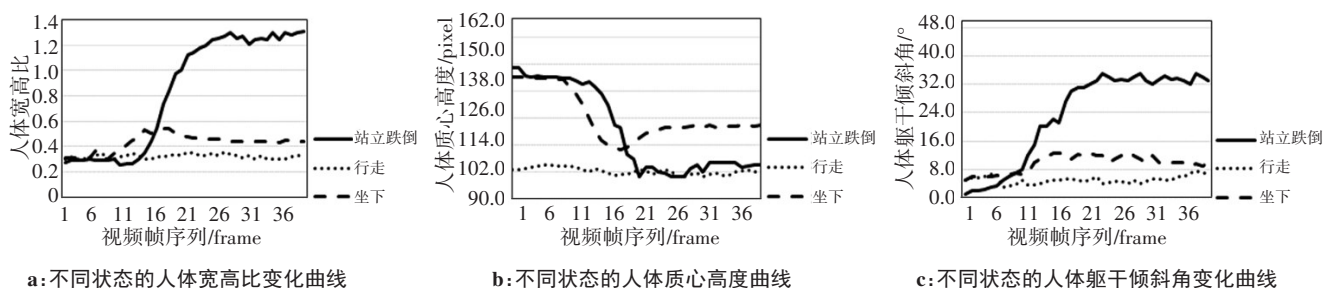


图5 场景信息变化图

Figure 5 Changes of scene information

对3D场景信息进行Sigmoid函数处理,避免不同的向量过大或过小,干扰分类准确度。将初步行为类别进行独热编码并与其置信度相乘得到行为类别向量,将编码后的行为类别向量与场景信息向量进行拼接经过全连接层,得到最终的跌倒检测结果。

2 实验结果与分析

实验过程在GPU服务器上完成,操作系统为Ubuntu 18.04.5,显卡为RTX2080Ti (11GBx1),软件平台为OpenCV4.5.3、Python3.8、CUDA11.0、PyTorch1.7.1。将原始图片尺寸为320×240缩放为224×224的尺寸进行训练,批次大小设置为16,迭代次数为200,采用Adam作为优化器,动量设为0.9,并使用动态学习率调整,初始学习率为1e-3。

2.1 模型评价指标

评价指标使用准确率(Accuracy, Acc)、精确率(Precision, Pr)、召回率(Recall, Re)作为模型评价指标,具体计算公式如下所示:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \quad (8)$$

$$\text{Pr} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (9)$$

$$\text{Re} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (10)$$

其中,TP为正确检测出的跌倒数量,TN为正确检测的非跌倒数量,FP非跌倒检测为跌倒的数量,FN是跌倒检测为非跌倒的数量。

2.2 数据集实验结果

为测试所提算法的可行性,使用公开的跌倒检测数据集URFD,数据集包含30个跌倒视频和40个日常行为视频,由两个不同的视角拍摄得到。实验

选用平视角度拍摄的视频序列,并对数据集进行扩充,使用intel D435相机对6位志愿者进行跌倒行为采集。扩充后的数据集共分为行走、站立跌倒、坐下、坐立跌倒、弯腰5类行为,每类包含60个视频,每个视频包含40~50张视频帧,图片分辨率为320×240。并按8:2的比例进行训练集和验证集的划分。实验利用Labelling工具手动标注目标框和动作类别。图6分别展示了站立跌倒和坐立跌倒两类跌倒行为。本文算法对于两种不同的跌倒行为均具有较好的表现,可以较好地检测日常行为。

2.3 不同策略的消融实验

为验证算法各个模块对于跌倒检测的可行性,对算法模块进行拆解,进行消融实验。表2中Baseline为基础的跌倒检测网络,3D卷积分支为深度可分离的3D卷积,2D部分为YOLOv4-tiny的Darknet部分,特征融合部分为扩张卷积,目标框损失函数为CIoU损失函数。3d_dense表示3D深度可分离卷积进行密集连接;2d_mix_resblock表示在2D卷积分支中重新设计的残差卷积块;SIoU表示预测框损失函数为SIoU;Non_local表示在特征融合部分融入非局部注意力机制;场景信息表示融合场景信息。“√”表示使用该模块,“×”表示未使用。算法2、3、4、5、6分别为使用以上5种不同策略对跌倒检测网络进行不同程度的优化。从表中可以看出不同的策略对于算法的表现呈现正向优化。

算法2相较算法1精确率提高0.7%,表明密集连接有效地增加了时序特征的提取能力,提高分类精确率。算法3相较算法2,召回率增加1.7%,说明改进的残差卷积块提升了网络的特征提取能力,降低漏检率。算法4与算法3相比,准确率提高1.7%,表

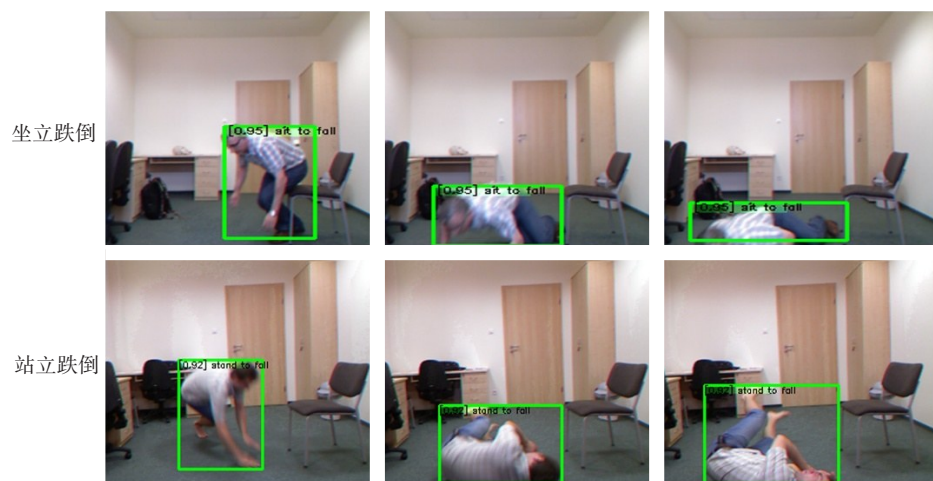


图6 数据集检测结果图

Figure 6 Test results on the dataset

表2 消融实验结果

Table 2 Ablation experiment results

方法	Baseline	3d_dense	2d_mix_reblock	SIoU	Non_local	场景信息	准确率/%	精确率/%	召回率/%
1	√	×	×	×	×	×	92.6	90.3	92.1
2	√	√	×	×	×	×	93.2	91.0	93.4
3	√	√	√	×	×	×	94.0	92.2	95.1
4	√	√	√	√	×	×	95.7	94.6	96.3
5	√	√	√	√	√	×	97.3	96.5	96.7
6	√	√	√	√	√	√	98.3	97.1	96.8

明 SIoU 使得算法更快地收敛,有利于提高算法准确度。算法 5 引入非局部注意力机制,增强特征融合能力,相较算法 4,准确率提高 1.6%,表明非局部注意力机制在特征融合部分更适合于跌倒检测。在算法的最后融入场景信息,能对类跌倒行为进行更好的区分,准确率相比未融合场景信息的算法提高 1.0%。

2.4 不同帧数的消融实验

分类准确度依赖输入 3D 卷积分支的连续帧,当输入的连续帧的数目和采样频率不同时,跌倒算法的准确率也受到不同的影响。为此,分别对不同数目和采样频率进行消融实验,结果见表 3。数目 f 为 8、16、32;采样频率 l 为 0、1, $l=0$ 表示无间隔采样, $l=1$ 表示每隔一帧间隔采样。当 $l=0$ 时,连续帧的数目与准确率成正相关,因为更多的连续帧包含更多的时序信息。当 $l=1$ 时,数目帧为 16 时,准确率最高,因为当采样频率为 1 时,数目为 8 时,连续帧含有的时序特征不够完整,当数目为 32 时,连续帧可能会包含跌倒动作以外不相关的动作帧,阻碍正确的跌倒动作帧的获取。

2.5 不同算法的对比实验

为验证算法的可行性,选取扩充后的 URFD 数据

表3 不同帧数与采样频率消融实验(%)

Table 3 Results of ablation experiment with different frame numbers and sampling frequency (%)

帧数	准确率	精确率	召回率
8($l=0$)	96.8	95.9	96.5
16($l=0$)	97.6	96.5	97.1
32($l=0$)	97.7	96.7	97.0
8($l=1$)	97.1	96.8	97.2
16($l=1$)	98.3	97.1	96.8
32($l=1$)	95.9	94.0	95.2

集的跌倒片段进行不同算法的对比实验。文献[22]首先使用姿态估计算法提取人员骨骼关键点信息,再使用长短期记忆网络进行行为分类,进行跌倒与非跌倒的分类;文献[23]提出一种多任务沙漏卷积自编码跌倒检测方法,通过视频帧重建增强网络的特征表现力;文献[24]提出一种基于关键点提取和 ST-GCN^[25]模型的实时跌倒检测算法;文献[26]基于光流法和卷积神经网络,进行室内跌倒检测。对比

实验选取坐立跌倒视频进行实验,实验结果见图 7。图中红色圆形为未检测到身体关键点标注,文献[24]脚踝关键点未正确检测。红色椭圆为误检标注,文献[22]在第 48、49、53 帧有 3 帧误检,文献[26]在第 51 帧有 1 帧误检,造成检测结果的不稳定。文献[24]红色框为跌倒,绿色框为非跌倒,其中在第 49、51、56 帧有 3 帧误检,误检较高是由于一阶段的关

键点检测中,脚踝关键点未检测出,导致 ST-GCN 模型不能正确判断跌倒行为。相比文献[22, 24, 26],文献[23]对于跌倒检测具有较好的表现,但是其不能对跌倒人员进行空间定位,且算法较复杂。本文算法绿色框为对于跌倒人员进行空间定位,能快速定位跌倒人员进行救治。

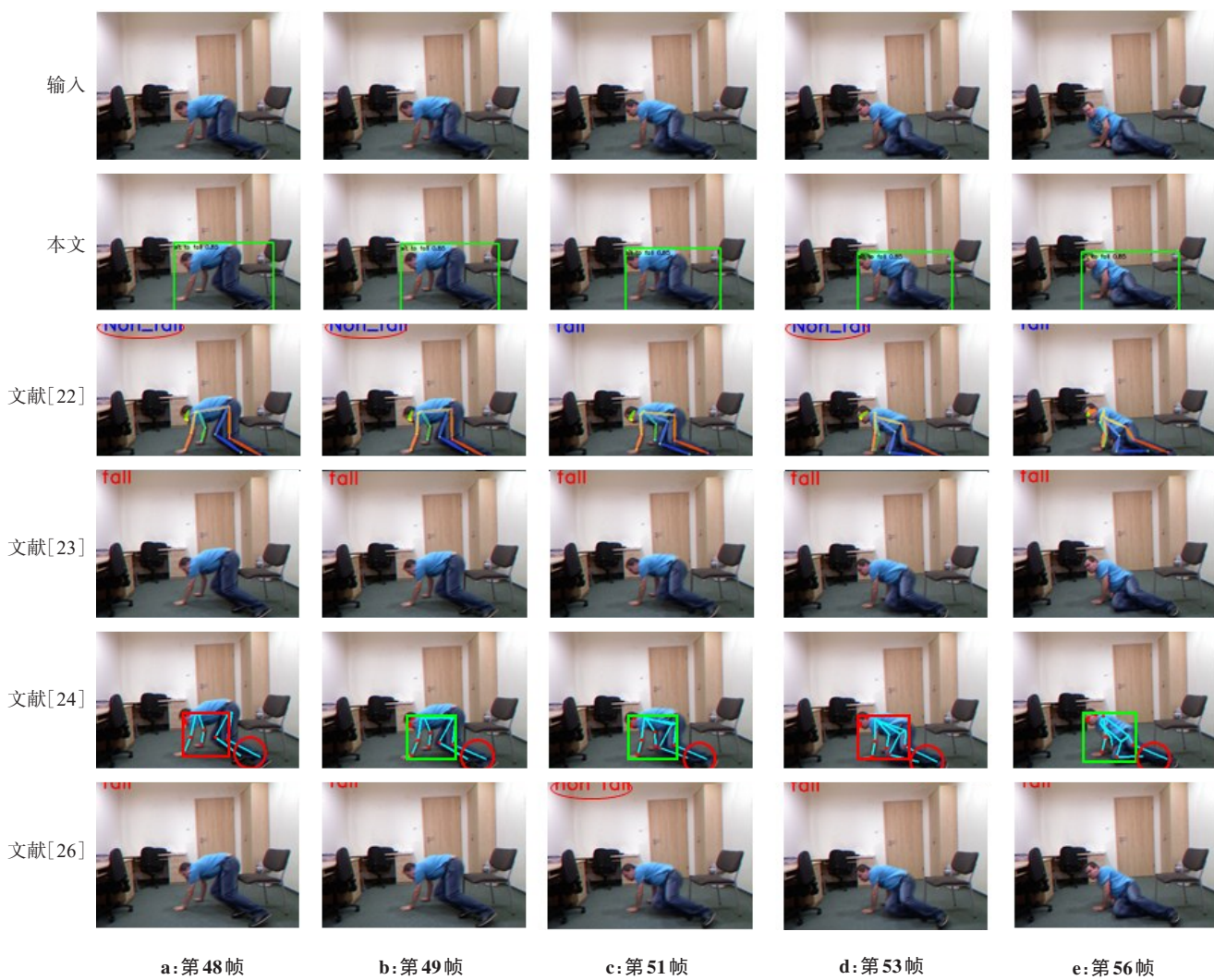


图 7 不同算法对比实验结果
Figure 7 Comparison of experimental results of different algorithms

模型评价指标对比见表 4,文献[22]和文献[24]为双阶段的跌倒检测方法,姿态估计和关键点检测的准确率较低时,使得后续的跌倒判断的准确率较低。文献[23]和文献[26]分别使用重建帧和双流法的方法,算法较复杂,且准确率低于本文算法。从表 4 可以看出,本文算法具有较高的准确率。

3 结 语

为更好地保障社区医疗养老健康,对跌倒的老

人进行及时发现并救治,本研究提出一种融合时序特征和空间特征的跌倒检测算法,算法的 3D 卷积分支融入密集连接从连续帧中提取时序信息,提高分类精确度;对于漏检的问题,改进 2D 卷积主干网络的残差块,从单帧图像中提取空间信息进行人员定位。在特征融合部分引入非局部注意力机制,并使用 SIoU 作为目标框的损失函数,提高检测框的定位精度。在算法的最后融合场景信息,3 种场景信息较好地区别于弯腰等类似跌倒行为。为验证算法性能,

表4 不同算法的对比试验结果(%)

Table 4 Results of comparative experiments with different algorithms (%)

方法	准确率	精确率	召回率
文献[22]	95.1	95.7	95.0
文献[23]	96.2	92.3	93.0
文献[24]	94.7	91.5	94.8
文献[26]	93.6	92.6	96.4
本文算法	98.3	97.1	96.8

在扩充后的URFD数据集进行对比实验。实验结果表明本文算法的准确率为98.3%,精确率为97.1%,召回率为96.8%。本文算法对于跌倒行为具有较好的表现,可以及时发现跌倒的老人,为独居老人的医疗健康提供保障。

【参考文献】

[1] 郝莲,李胜,向芳,等. 2003-2018年中国跌倒死亡率的时间趋势分析[J]. 现代预防医学, 2020, 47(23): 4357-4360.
Hao L, Li S, Xiang F, et al. Time trend of falls mortality in China, 2003-2018[J]. Modern Preventive Medicine, 2020, 47(23): 4357-4360.

[2] 任晶秋,蒋杨,张光华. AHRs的老人跌倒检测算法[J]. 电子测量与仪器学报, 2020, 34(12): 190-196.
Ren JQ, Jiang Y, Zhang GH. The algorithm of elderly fall detection based on AHRs [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(12): 190-196.

[3] 王晓雷,李栋豪,郑晓婉,等. 基于RBF神经网络的跌倒检测算法研究[J]. 电子测量与仪器学报, 2019, 33(11): 185-191.
Wang XL, Li DH, Zheng XW, et al. Research on fall detection algorithm based on RBF neural network[J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(11): 185-191.

[4] 陈涵,余磊,彭泗田,等. 基于红外图像和逆向投影算法的室内人体跌倒检测方法[J]. 红外技术, 2021, 43(10): 968-978.
Chen H, Yu L, Peng ST, et al. Indoor human fall detection method based on infrared images and back-projection algorithm[J]. Infrared Technology, 2021, 43(10): 968-978.

[5] 徐甲栋,陈强,徐一雄,等. 基于MEMS传感器的实时跌倒检测系统设计[J]. 传感器与微系统, 2022, 41(7): 77-80.
Xu JD, Chen Q, Xu YX, et al. Design of real time fall detection system based on MEMS sensor [J]. Transducer and Microsystem Technologies, 2022, 41(7): 77-80.

[6] 曹建荣,吕俊杰,武欣莹,等. 融合运动特征和深度学习的跌倒检测算法[J]. 计算机应用, 2021, 41(2): 583-589.
Cao JR, Lü JJ, Wu XY, et al. Fall detection algorithm integrating motion features and deep learning [J]. Journal of Computer Applications, 2021, 41(2): 583-589.

[7] Redmon J, Farhadi A. YOLOv3: an incremental improvement[J]. arXiv: 1804.02767, 2018.

[8] 马露,裴伟,朱永英,等. 基于深度学习的跌倒行为识别[J]. 计算机科学, 2019, 46(9): 106-112.
Ma L, Pei W, Zhu YY, et al. Fall action recognition based on deep

learning[J]. Computer Science, 2019, 46(9): 106-112.

[9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector [J]. arXiv: 1512.02325, 2015.

[10] Tran D, Bourdev L, Fergus R, et al. Learning spatiotemporal features with 3D convolutional networks[J]. arXiv: 1412. 0767, 2014.

[11] 李秀智,张冉,贾松敏. 面向助老行为识别的三维卷积神经网络设计[J]. 北京工业大学学报, 2021, 47(6): 589-597.
Li XZ, Zhang R, Jia SM. Design of 3D convolutional neural network for action recognition for helping the aged[J]. Journal of Beijing University of Technology, 2021, 47(6): 589-597.

[12] Li SW, Man CH, Shen A, et al. A fall detection network by 2D/3D spatio-temporal joint models with tensors compression on edge[J]. ACM Trans Embed Comput Syst, 2022, 21(6): 1-19.

[13] 马敬奇,雷欢,陈敏翼. 基于Alphapose优化模型的老人跌倒行为检测算法[J]. 计算机应用, 2022, 42(1): 294-301.
Ma JQ, Lei H, Chen MY. Fall behavior detection algorithm for the elderly based on Alphapose optimization model [J]. Journal of Computer Applications, 2021, 42(1): 294-301.

[14] Fang HS, Xie SQ, Tai YW, et al. RMPE: regional multi-person pose estimation[C]//2017 IEEE International Conference on Computer Vision. IEEE, 2017: 2353-2362.

[15] 孙颖,张吟龙,王鑫,等. 面向医疗护理的视觉监控医院患者跌倒检测[J]. 中国医学物理学杂志, 2022, 39(4): 436-441.
Sun Y, Zhang YL, Wang X, et al. Medical care oriented visual surveillance of patient falls in the hospital[J]. Chinese Journal of Medical Physics, 2022, 39(4): 436-441.

[16] Cao Z, Simon T, Wei S, et al. Realtime multi-person 2D pose estimation using affinity fields [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 1302-1310.

[17] Wang XL, Girshick R, Gupta A, et al. Non-local neural networks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 7794-7803.

[18] Gevorgyan Z. SIoU loss: more powerful learning for bounding box regression[J]. arXiv: 2205.12740, 2022.

[19] Huang G, Liu Z, Van L, et al. Densely connected convolutional networks [C]//IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 4700-4708.

[20] Bochkovshiy A, Wang C, Liao H. YOLOv4: optimal speed and accuracy of object detection[J]. arXiv: 2004.10934, 2020.

[21] Zhang XY, Zhou XY, Liu MX, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [J]. arXiv: 1707.01083, 2017.

[22] Salimi M, Machado J, Tavares J. Using deep neural networks for human fall detection based on pose estimation[J]. Sensors, 2022, 22: 4544-4559.

[23] Cai X, Li SY, Liu XY, et al. Vision-based fall detection with multi-task hourglass convolutional auto-encoder [J]. IEEE Access, 2020, 8: 44493-44502.

[24] 王君涛,潘长乐,杨龙飞,等. 基于改进的ST-GCN模型的跌倒检测算法[J]. 信息技术与信息化, 2022, 2: 69-71.
Wang JT, Pan CL, Yang LF, et al. Fall detection algorithm based on an improved ST-GCN model [J]. Information Technology and Informatization, 2022, 2: 69-71.

[25] Yan SJ, Xiong YJ, Lin DH. Spatial temporal graph convolutional networks for skeleton-based action recognition[J]. arXiv: 1801.07455, 2018.

[26] 孙博文,祁燕,杨大为,等. 基于光流法和卷积神经网络的跌倒检测方法[J]. 微机处理, 2021, 42(2): 45-48.
Sun BW, Qi Y, Yang DW, et al. Falling behavior detection method based on optical flow and convolutional neural network [J]. Microprocessors, 2021, 42(2): 45-48.

(编辑:谭斯允)