

文章编号: 2095-2163(2023)11-0298-06

中图分类号: TP18

文献标志码: A

# 人体摔倒行为识别研究中的自动化深度特征提取

潘道华<sup>1,2</sup>, 刘宏伟<sup>1</sup>

(1 哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001; 2 黑龙江民族职业学院 电子信息工程系, 哈尔滨 150001)

**摘要:** 在人体摔倒行为识别的研究中, 特征提取是非常重要的一个环节。目前, 特征提取主要是通过手工法提取, 提取的特征包括时域特征、频域特征和时频域特征。这些传统的手工法提取的特征往往导致分类识别的准确性不够高。为了提高识别准确性和特征提取的自动化程度, 减少人为的干预, 本文提出用基于卷积神经网络(CNN)进行深度特征的提取, 然后再经过机器学习分类器算法进行识别, 利用该方法可以大大提高分类识别的准确率。

**关键词:** 特征提取; 摔倒行为识别; CNN; 深度特征

## Automatic depth feature extraction in human falling recognition research

PAN Daohua<sup>1,2</sup>, LIU Hongwei<sup>1</sup>

(1 School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China;

2 Department of Electronic and Information Engineering, Heilongjiang Minzu College, Harbin 150001, China)

**Abstract:** Feature extraction is a very important link in the study of human falling behavior recognition. At present, feature extraction is mainly through manual method, and extracted features includes time-domain features, frequency-domain features and time-frequency-domain features. The traditional manual extraction methods often results in a low accuracy of classification and recognition. In order to improve the accuracy of recognition and the automation of feature extraction, also in order to reduce human intervention, this paper proposes a method based on Convolutional Neural Network (CNN) to extract deep features and identify the human activities through machine learning classifier, which can greatly improve the accuracy of classification and recognition.

**Key words:** feature extraction; recognition of falling behavior; CNN; depth feature

## 0 引言

当今社会人口老龄化现象是最显著的人口特征之一, 由于老年人生理结构的老化和身体机能的下降, 在生活中老年人意外摔倒不再是个例, 是造成老年人严重健康伤害、甚至死亡的十大主要原因之一。在日常生活中, 尤其是在潮湿处, 如厨房、浴室等地, 老人就容易出现打滑。此时, 若老年人还患有脑溢血、心脏病等疾病, 发生摔倒就有可能危及老年人的生命。此外, 当今社会老人户外摔倒, 若未能遇到路人及时伸出援手, 就将导致摔倒后的老人很长一段时间内得不到救助<sup>[1]</sup>, 这种情况就更会加剧对老年人的伤害。综上可知, 老年人摔倒问题已经衍生成为一个重大的社会问题<sup>[2]</sup>。因此, 摔倒行为识别对老年人尤为重要, 有关人体摔倒行为识别的研究具有很高的实际应用价值。

在人体摔倒行为识别的研究中, 特征提取是至关重要的一个环节。目前, 特征提取主要是通过手工法提取, 提取的特征包括时域特征、频域特征和时

频域特征。这些传统的手工法提取的特征往往导致分类识别的准确性不够高。为了提高识别准确性和特征提取的自动化程度, 减少人为的干预, 本文提出基于卷积神经网络(CNN)进行深度特征的提取, 而后再经过机器学习分类器算法进行识别的方法。将深度学习技术与机器学习算法相相合, 可以使整个模型的综合性能得以提高, 从而提升人体摔倒行为识别准确率。将 CNN 用作分类器会有一定的局限性。分析可知, CNN 是多层感知机的拓展模型, 通过使用经验风险最小化的算法来进行训练学习, 这就容易导致模型最终得到的不是全局最优解、而是局部最优解<sup>[3]</sup>。另外, 当数据到达网络的输出层(output layer)时, 经常使用的多分类激活函数(*softmax*)往往倾向于把一个极大的值(近似为 1)赋给某一个类别, 而其它的类别所得的值相比就很小(近似为 0), 这就明显降低了 CNN 在分类识别中的容错能力。因此, CNN 并不适合用作分类器模型<sup>[4]</sup>。本文选择了 CNN 结合 SVM 的混合式模型做人体摔倒行为的分类研究, 可以大大提高识别的准确率。

**基金项目:** 黑龙江省教育科学规划 2023 年度重点课题(ZJB1423035)。

**作者简介:** 刘宏伟(1971-), 男, 博士, 教授, 主要研究方向: 并行计算与体系结构、容错计算机、云计算与资源调度。

**通讯作者:** 潘道华(1981-), 女, 博士研究生, 主要研究方向: AI、模式识别、可穿戴计算。Email: pandaohua@fctl.hit.edu.cn。

收稿日期: 2021-03-20

## 1 人体摔倒过程的分析及数据获取

### 1.1 人体摔倒过程的分析

人体摔倒在地面或接近于地面的较低物体上时,会对其产生一定的冲击,根据冲击的强度,可以将摔倒分为硬摔倒和软摔倒两大类<sup>[5]</sup>。其中,硬摔倒是指发生撞击后,会引起身体的一系列疾病的摔倒。软摔倒是指因为撞击强度低,发生撞击后,身体不会受到严重伤害的摔倒,比如老年人靠在墙边,或者扶着墙壁等比较缓慢地摔倒。由于硬摔倒会对身体产生较大的危害,因此,本文着重对这种摔倒进行研究和分析。根据人体动力学的原理,当人体的重心投向地面的投影不在人体的支撑点上时,人体的力平衡机制就已经被打破。如果人体的重心或支撑点没有得到及时调整,就极易发生摔倒。这种人体的不平衡状态导致的结果是人体以一定的加速度向下跌落,同时,身体的姿态角度发生变化。根据这一分析,可以将摔倒过程分为 4 个阶段:

(1) 摔倒开始阶段:此时摔倒并没有发生,但人体平衡被破坏了,在这一阶段里身体加速度和身体姿态角度都没有发生改变。

(2) 身体下降阶段:从人体失去平衡开始,到身体与地面接触结束。在这个过程中,人的双脚着地面会逐渐脱离地面,身体在重力加速度的作用下发生向下的坠落,该阶段中,人体处于失重的状态。由于人体受到重力加速度的作用,身体的变化速度在与地面发生撞击之前会逐渐增大,直至达到某个最大值。身体的加速度值也会发生变化,由重力加速度  $G$  逐渐下降为接近于 0。身体绕人体坐标系中的坐标轴旋转的旋转角度和旋转角速度发生变化。

(3) 冲击发生阶段:从身体触碰地面或者较低平面那一刻开始,到身体静止在地面上为止。在冲击阶段中,身体的加速度值会发生剧烈的变化,身体将承受很大的冲击,这一阶段也是造成身体损伤的阶段。在身体受到撞击发生前,身体向下的运动速度达到最大值,此后与地面或其它较低平面发生突发性碰撞,速度发生骤减,减至 0。此时,加速度会迅速地变大,会形成一个明显的峰值。

(4) 身体静止阶段:身体在受到冲击后通常会躺在地面上一段时间。一般情况下,无论身体是否受到伤害,在摔倒后的一段较短的时间内,身体都会处于一种相对静止的状态。在这一阶段,身体的加速度值处于相对平稳的状态。

### 1.2 数据获取

本文的实验数据是通过志愿者穿戴着一套动作捕捉设备进行采集的。该套设备全身共包含 16 个传感器节点,如图 1 所示,每个传感器节点都为九轴传感器,分别集成了三轴加速度计、三轴陀螺仪、三轴地磁计。

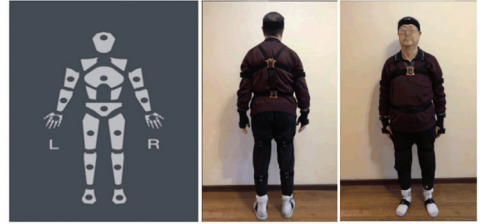


图 1 数据采集所用的传感器节点

Fig. 1 The sensor nodes used for data acquisition

该套动捕设备通过 WiFi 与主机通信,并将采集到的传感器数据存储在主机中。由于本文的实验是做人体大幅度动作的分类,旨在有效识别出是否有摔倒行为的发生,所以手背处的传感器节点所采集的数据对于摔倒识别的影响不大。另外,由于全身的传感器呈身体对称性分布,所以本文实验只选用身体单侧肢体的传感器节点和身体主干上的传感器节点的数据就可以实现实验要求,文中选用了 7 个关键的传感器节点所采集的传感器数据进行实验,这 7 个传感器节点分别位于头部、胸部、腰部、上臂部、前臂部、大腿部和小腿部。

本文的部分实验数据是来源于 5 位 65 岁以上的老年人真实日常生活中的各种行为数据的采集,但是由于没有那么多老年人可以进行实验数据的采集,另外又考虑到老年人的安全和健康,本文的另一部分数据获取来源于征集的 20 位青年志愿者,志愿者的年龄范围为 18~22 岁,平均年龄为 20.5 岁;体重范围为 46~79 kg,平均体重为 62.18 kg;身高范围为 155~178 cm,平均身高为 169.6 cm。所有的志愿者模拟日常生活中 7 类常见的行为活动<sup>[6]</sup>:行走、坐、蹲、躺、站立、上楼、下楼,每种日常行为活动的数据采集持续时间为 5 min。每位志愿者进行了模拟摔倒,摔倒分为向前、向后、向左、向右和行走中摔倒(也称为绊倒)。分别模拟人体下肢在突然失去控制的情况下和发生眩晕的情况下发生的摔倒,每种摔倒进行 20 次,共计 100 次,所有的摔倒都是在厚约 10 cm 的海绵垫上进行的。所有传感器节点的采样频率均为 120 Hz。

实验采集的原始传感器数据信号通过 0.2 Hz 的三阶巴特沃斯高通滤波器<sup>[7]</sup>进行滤波,接着再通过滑动窗口技术实现传感器数据的分割,将数据分

割成若干个数据片段。滑动窗口的时间窗口大小是由目标行为活动集来决定的,最终根据实际情况,将滑动窗口的时间设定为 3 s,这是一个比较常用的窗口大小<sup>[8]</sup>。7 个传感器节点分别穿戴在人体上的 7 个不同位置,因为每个传感节点都包括三轴加速度计、三轴陀螺仪、三轴地磁计,所以每个数据文件包括 63 列传感器数据。

## 2 特征提取

### 2.1 传统特征的提取

特征提取是进行人体摔倒行为识别的过程中不可缺少的主要环节,对摔倒行为识别的准确性有着很大的影响。传统的统计特征会经常被用在人体行为识别的研究中,通常可分为时域特征、频域特征和时频域特征。Khan 等人<sup>[9]</sup>和 Yan 等人<sup>[10]</sup>对时域特征和频域特征进行了对比,结果是在大多数情况下时域特征对分类的作用更大,但是将两者结合起来是一种更好的方法。

本文进行了传统特征的提取。对传感器的每个轴所采集的数据分别提取了 9 种常用特征:均值、方差、中位数、均方根、标准差、最大差值、偏度、峰度和能量,各指标的数学定义公式如下。

(1) 均值。计算公式可写为:

$$\mu = \frac{\sum_{i=1}^N T(i)}{N} \quad (1)$$

(2) 方差。计算公式可写为:

$$D = \sum_{i=1}^N (T(i) - \mu)^2 \quad (2)$$

(3) 中位数。计算公式可写为:

$$MED = \text{mediam}(T(i)) \quad (3)$$

(4) 均方根。计算公式可写为:

$$R = \sqrt{\frac{\sum_{i=1}^N T(i)^2}{N}} \quad (4)$$

(5) 标准差。计算公式可写为:

$$\sigma = \sqrt{\frac{D}{N-1}} \quad (5)$$

(6) 最大差值。计算公式可写为:

$$\Delta T = \max(T(i)) - \min(T(i)) \quad (6)$$

(7) 偏度。计算公式可写为:

$$S = \frac{1}{N} \sum_{i=1}^N \left[ \frac{T(i) - \mu}{\sigma} \right]^3 \quad (7)$$

(8) 峰度。计算公式可写为:

$$K = \frac{1}{N} \sum_{i=1}^N \left[ \frac{T(i) - \mu}{\sigma} \right]^4 - 3 \quad (8)$$

(9) 能量。计算公式可写为:

$$E = \sqrt{\frac{\sum_{i=1}^N F(i)^2}{N}} \quad (9)$$

其中,  $T(i)$  表示经过预处理后的数据的第  $i$  个采样点值,  $F(i)$  为  $T$  经过傅立叶变换后的第  $i$  个傅立叶系数。

经过特征提取后,形成了由各个传感器数据片段的特征向量组成的多维特征向量。

### 2.2 深度特征的提取

本文使用卷积神经网络 CNN 来进行深度特征的提取。为了提取深度特征,构建 CNN 的结构<sup>[11]</sup>如图 2 所示。在实验部分,对基于深度特征的摔倒行为识别效果和基于传统的统计特征的识别效果做了对比。

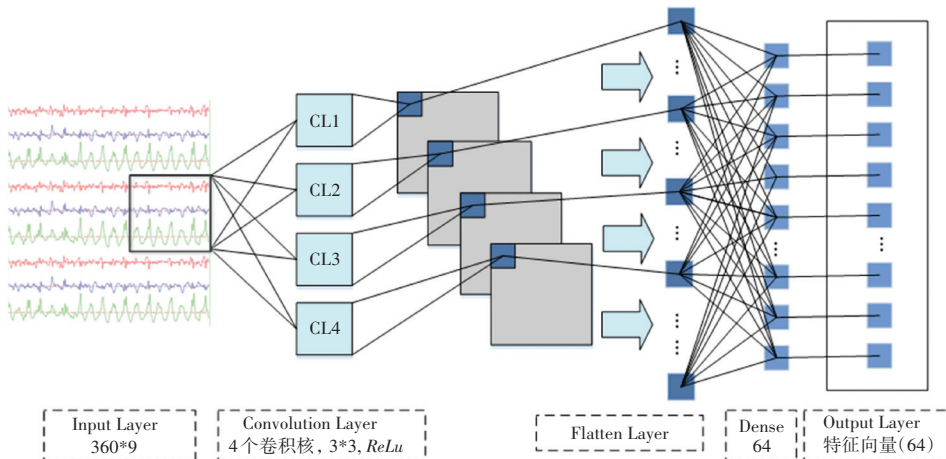


图2 进行深度特征提取的 CNN 结构

Fig. 2 CNN structure for depth feature extraction



为了阐述模型的工作机制, 用  $X$  来表示卷积神经网络的输入数据。经过卷积后, 第  $k$  个特征图中的元素  $E_{p,q}^k$  被记为<sup>[12]</sup>:

$$E_{p,q}^k = F(b^k + \sum_{\alpha=1}^i \sum_{\beta=1}^j \omega_{\alpha,\beta}^k X_{\alpha,\beta}^{p*(m-2)+q}) \quad (10)$$

其中,  $(p, q)$  表示元素  $E_{p,q}^k$  在对应的特征图中的具体位置;  $b^k$  是偏置项; 卷积核的大小用  $i * j$  来表示;  $\omega_{\alpha,\beta}^k$  表示卷积核权重值, 这里的  $\alpha$  和  $\beta$  标识为该权重在卷积核中的位置。

假设原始传感器数据序列大小为  $m * n$ ,  $S_{\alpha,\beta}^{p*(m-2)+q}$  表示进行第  $p * (m - 2) + q$  次卷积运算时输入的数据中对应位置的值。  $F$  为卷积层的激活函数, 本文使用的是 *ReLU* 线性整流函数, 函数形式可表示为:

$$F(x) = \max(0, x) \quad (11)$$

经过平化层后, 所有的特征图经由矩阵转换为 1D 的向量。此后数据再经过全连接层, 输出深层特征。全连接层的数学原理公式可表示为:

$$v_\gamma = S(\sum_{\gamma=1}^u \mathbf{W} x_\gamma + \mathbf{b}) \quad (12)$$

其中,  $u$  表示输入的数据的维度;  $x_\gamma$  个输入的元素;  $v_\gamma$  表示第  $\gamma$  个输出的元素;  $\mathbf{W}$  表示权重向量;  $\mathbf{b}$  为偏置值;  $S$  是激活函数  $\tanh$ , 其值可由如下数学公式计算求出:

$$S(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (13)$$

最后一层的输出向量  $\mathbf{V} = [v_1, \dots, v_u]$  是卷积神经网络提取出的深度特征, 为下一步将要进行的机器学习分类器做训练使用的特征。

每个节点的采集数据是由 9 个轴的数据组成, 采集到的数据序列经过滑动窗口进行分割, 分割为每个窗口包含 360 个采样点的传感器数据文件。因此, 进行深度特征提取的输入为多个  $360 * 9$  大小的矩阵。卷积层采用 4 个卷积核, 因为每 3 个传感器数据都属于同一类型, 所以使用了 4 个大小为  $3 * 3$  的卷积核。将 *stride* 设置为 1。经过卷积运算后的结果, 再经过平坦层的一维化运算进入稠密层, 该层中设计了 64 个权重节点, 经过全连接运算后得到 64 维的特征向量。

### 3 实验对比分析

人体摔倒行为的识别一般采用敏感性、特异性、准确性三个指标来评价。这 3 个指标是用 4 种可能情况来计算的, 参见表 1。

表 1 摔倒行为识别的 4 种可能情况

Tab. 1 Four possible cases of falling behavior recognition

	摔倒未发生	摔倒发生
没有检测到摔倒	真阴性 (TN)	假阴性 (FN)
检测到摔倒	假阳性 (FP)	真阳性 (TP)

表 1 中, True Negatives (TN) 表示摔倒并未发生, 可以正确识别为不是摔倒、而是其它行为的样本数; False Negatives (FN) 表示摔倒发生了、但却没有被识别出来的摔倒行为的样本数; False Positives (FP) 表示摔倒并未发生、但却把其它的日常行为活动误判为是摔倒行为的样本数; True Positives (TP) 表示摔倒发生了、而且能正确地识别出来的摔倒行为的样本数。

进而, 针对研究采用的评价指标, 拟做剖析分述如下。

(1) 敏感性。是标识检测摔倒能力的指标。可由摔倒发生并能够成功检测到的摔倒数 (TP) 与发生的总摔倒数 (TP + FN) 之比计算得出:

$$Sensitivity = \frac{TP}{TP + FN} \quad (14)$$

(2) 特异性。是标识避免误报能力的指标, 计算公式为:

$$Specificity = \frac{TN}{TN + FP} \quad (15)$$

至此分析可知, 敏感性和特异性用来更好地了解摔倒行为识别的某些限制。

(3) 准确性。是有效区分和识别摔倒 (TP) 和非摔倒 (TN) 的能力, 可以通过以下公式计算得出:

$$Accuracy = \frac{TP + TN}{P + N} \quad (16)$$

其中,  $P$  表示摔倒发生的次数,  $N$  表示没有摔倒发生的次数。

总体来说, 准确性是一个全局性的指标, 本文实验部分把准确性作为摔倒行为识别的衡量标准。

#### 3.1 基于传统特征的摔倒识别

通过 2.1 节提取的统计特征作为特征向量, 利用支持向量机 (SVM)<sup>[13]</sup>、决策树 (DT)<sup>[14]</sup>、随机森林 (RF)<sup>[15]</sup> 三种分类器模型分别进行训练与测试, 然后采用十折交叉验证法对分类效果进行验证。准确性的实验比较结果见表 2。

表 2 显示了 3 种分类器算法的不同检测效果。可能是因为 DT 分类器和 RF 分类器不适合对连续数据进行处理, 还可能因为对分类器参数的设置没有达到最优, 所以导致 RF 的效果略差, DT 的效果

更差。总的说来,SVM 的平均识别准确性最高,整体效果最好。

表 2 基于统计特征的摔倒行为识别准确性

Tab. 2 Accuracy of falling behavior recognition based on statistical features

传感器位置	SVM	DT	RF
头部	85.70%	75.88%	84.19%
胸部	85.18%	69.53%	83.86%
腰部	89.61%	86.46%	88.23%
前臂部	81.62%	69.93%	80.49%
上臂部	82.30%	73.53%	81.70%
小腿部	82.71%	75.78%	82.92%
大腿部	85.42%	77.17%	84.22%

另外,传感器佩戴的不同位置对摔倒行为的检测准确率有一定的差别。腰部采集的实验数据效果最好。不同传感器佩戴位置、不同的分类器算法对摔倒行为的检测准确率有一定的影响,但差别并不明显。综合 SVM、DT、RF 三种分类器算法的摔倒行为识别准确率对比效果,如图 3 所示。

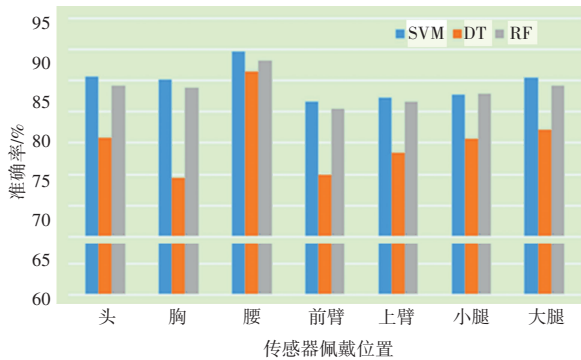


图 3 传感器佩戴位置与摔倒行为识别准确性的关系

Fig. 3 The relationship between sensor position and the accuracy of falling behavior recognition

通过实验可以看出:不同的传感器佩戴位置对摔倒行为的识别准确率影响并不大,特别是对 SVM 这个整体性能比较好的分类器。虽然方法不同,但这一结论与其他研究工作者得出的结论相似<sup>[16]</sup>。

### 3.2 基于深度特征的摔倒识别

#### 3.2.1 基于深度特征的摔倒行为识别的效果

利用本文设计的 CNN 结构提取新的深度特征,构成特征向量,利用支持向量机(SVM)、决策树(DT)、随机森林(RF)三种分类器模型分别对身体的 7 个传感器节点采集到的数据进行训练和测试,基于统计特征对身体各传感器节点的摔倒行为识别的准确性实验结果见表 3。

表 3 基于深度特征的摔倒行为识别准确性

Tab. 3 Accuracy of falling behavior recognition based on depth feature

传感器位置	SVM	DT	RF
头部	92.15%	77.83%	87.41%
胸部	90.77%	71.40%	83.11%
腰部	94.89%	88.05%	90.31%
前臂部	84.52%	73.82%	82.51%
上臂部	90.92%	75.20%	85.94%
小腿部	86.47%	77.92%	83.27%
大腿部	88.71%	78.36%	83.05%

#### 3.2.2 基于深度特征与传统特征的摔倒行为识别的效果对比

为了便于清楚地对比基于深度特征与传统特征的摔倒行为识别的效果,将识别准确率的效果以图 4 显示。

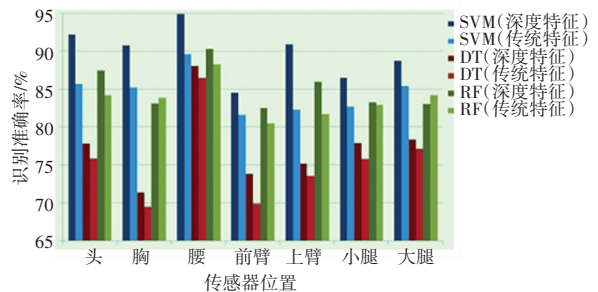


图 4 使用深度特征与传统特征进行的摔倒行为识别准确性的对比图  
Fig. 4 Falling behavior identification accuracy using depth features versus traditional features

由图 4 可以看出基于深度特征的人体摔倒行为识别的准确率明显高于基于传统统计特征的人体摔倒行为识别的准确率。图 4 中,分类器算法 SVM 通过 CNN 提取的深度特征对摔倒识别准确率的提升最为明显。另外,在对人体摔倒行为识别的整个实验过程中,可以发现无论使用哪一种分类器,相对运动幅度比较小的活动的识别精度都相对较高,这是因为相对运动幅度比较小的活动所具有的平稳信号特性。而对于运动时幅度比较大的动态行为活动,识别准确率略低一些,也许是因为受数据失真的影响更大的结果。但是本文采用的基于 CNN 提取的深度特征的方法可以大大地缓解这个问题。

## 4 结束语

本文提出了基于 CNN 结构进行深度特征提取的方法,将深度特征的提取结合 SVM 分类器用于人体摔倒行为识别的实验研究中,结果表明,提取不同的特征对人体活动识别的准确性有着很大的影响,