

基于3D视觉分析技术的医院患者跌倒识别方案设计

许庆¹, 陈文涛², 刘志文²

(1. 广州医科大学附属肿瘤医院, 广东 广州 510000;

2. 武汉大学中南医院, 湖北 武汉 430000)

摘要:为实现针对医院患者跌倒行为的准确识别,研究提出了一套基于3D视觉分析技术的人体跌倒识别方案。详细介绍了该方案中人体运动目标检测与人体行为特征提取的实现策略,并通过深度森林算法对所采集到的样本数据加以处理,进而实现对于跌倒行为的判别,最后通过模拟实验的方式对该方案的有效性进行分析。

关键词:跌倒识别; 3D视觉; YOLOv3算法; 人体行为; 特征提取

中图分类号: TP391.413 文献标识码: A 文章编号: 1003-7241(2025)04-0099-04

Design of Fall Recognition Scheme for Hospital Patients Based on 3D Visual Analysis Technology

XU Qing¹, CHEN Wen-tao², LIU Zhi-wen²

(1. Cancer Hospital Affiliated to Guangzhou Medical University, Guangzhou 510000 China;

2. Zhongnan Hospital of Wuhan University, Wuhan 430000 China)

Abstract: In order to realize the accurate recognition of hospital patients' fall behavior, this study proposes a set of human fall recognition scheme based on 3D visual analysis technology, introduces in detail the implementation strategies of human motion target detection and human behavior feature extraction in this scheme, and processes the collected sample data through deep forest algorithm, so as to realize the discrimination of fall behavior. Finally, the effectiveness of the scheme is analyzed by simulation experiments.

Keywords: fall recognition; 3D visual; YOLOv3 algorithm; human behavior; feature extraction

0 引言

人体运动状态检测是我国医学界的一项重要研究课题,随着我国老龄化趋势的不断加剧,医疗机构需要通过更加先进的技术来实现对于老年患者跌倒行为的无接触识别。甘惠兴等人提出了一套基于激光与视觉图像传感器的人体状态识别方案^[1],该技术虽然能够较为准确地检测出被测者的行走状态,但同时也依赖造价较为高昂的可穿戴设备,技术成本相对较高;李纪鑫等人提出了一套基于机器视觉图像的人体运动状态识别方案^[2],能够较为准确地对运动员的形体状态加以还原,然而该方案只支持2D图像识别,并且对于原始图像的清晰度要求较高,难以适用于医院这类需要进行长时间、大规模检测的公共场所。因此,本次研究提出了一套基于3D视觉分析技术的医院患者跌倒识别方案,通过3D视觉分析技术来提取患者形体的人体行为特征,并通过深度森林算法对被测者的跌倒行为加以识别。

1 人体运动行为分析

1.1 人体运动目标检测

本次研究通过YOLOv3算法对人体运动目标进行检测,该算法是YOLO的第三个迭代版本,新版本以聚类的方式在数据集上添加了9个先验框,可实现针对特征图边界框的精准预测,预测原理如图1所示^[3-4]。

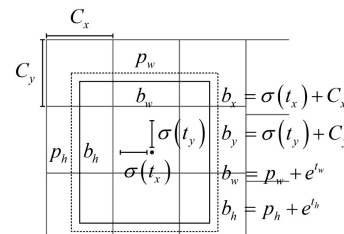


图1 YOLOv3边界框预测原理

YOLOv3算法能够计算出当前网格中心位置的偏移量,采用sigmoid函数对偏移量实施归一化处理,在提高预测精度的同时使模型更加稳定^[5-6]。在YOLOv3所输出的特征图中,各网格均带有3个边界框,各边界框均带有 t_x 、 t_y 、 t_w 、 t_h 、 t_o 五个参数值,其中 t_o 为置信度, t_w 和 t_h 为宽高度缩放, t_x 和 t_y 为预测框的中心偏移量。在图1中, p_h

和 p_w 分别为先验框的高和宽, C_x 和 C_y 用于描述当前网格相对于图像左上角的偏移量, 相应的预测值计算方法如下:

$$b_x = \sigma(t_x) + C_x \quad (1)$$

$$b_y = \sigma(t_y) + C_y \quad (2)$$

$$b_w = p_w + e^w \quad (3)$$

$$b_h = p_h + e^h \quad (4)$$

$$\sigma(t_o) = \Pr(\text{object}) * \text{IOU}(b, \text{object}) \quad (5)$$

损失函数设计是目标检测工作的一项重要任务, YOLOv3的损失函数由目标分类损失 L_{class} 、目标置信度损失 L_{conf} 和 L_{coord} 坐标损失三部分构成, 表达方式如下:

$$\text{loss}(\text{object}) = L_{\text{coord}} + L_{\text{conf}} + L_{\text{class}} \quad (6)$$

坐标损失 L_{coord} 负责对边界框坐标值的预测偏差加以约束, 其表达方式如下:

$$L_{\text{coord}} = \sum_i \sum_j^{K \times K} I_{ij}^{\text{obj}} \left[(-x_i * \log \hat{x}_i) - (1 - x_i) * \log(1 - \hat{x}_i) \right] + \sum_i \sum_j^{K \times K} I_{ij}^{\text{obj}} \left[(-y_i * \log \hat{y}_i) - (1 - y_i) * \log(1 - \hat{y}_i) \right] + \sum_i \sum_j^{K \times K} I_{ij}^{\text{obj}} \left[(w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2 \right] \quad (7)$$

式(7)中, M 代表单个网格所生成的边界框数量记为 M ; $K \times K$ 代表网格数量; I_{ij}^{obj} 用于判断第 i 个网格中第 j 个边界框是否包含目标, 若包含则记为1, 若不包含则记为0. x_i, y_i, w_i, h_i 代表标签值, $\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i$ 代表预测值.

置信度损失 L_{conf} 负责对由目标置信度所引起的误差加以约束, 本次研究所提到的置信度指的是边界框覆盖目标特征的准确度, 表达方式如下:

$$L_{\text{conf}} = \sum_i \sum_j^{K \times K} I_{ij}^{\text{obj}} \left[(-C_i * \log \hat{C}_i) - (1 - C_i) * \log(1 - \hat{C}_i) \right] + \sum_i \sum_j^{K \times K} I_{ij}^{\text{noobj}} \left[(-C_i * \log \hat{C}_i) - (1 - C_i) * \log(1 - \hat{C}_i) \right] \quad (8)$$

式(8)中, I_{ij}^{noobj} 用于判断第 i 个网格中第 j 个边界框是否包含目标, 若包含则记为0, 若不包含则记为1, C_i 和 \hat{C}_i 分别代表标签值和预测值.

目标类别损失 L_{class} 负责对目标类别的预测偏差加以约束, 该数值可通过二值交叉熵损失函数进行计算, 具体方式如下:

$$L_{\text{class}} = \sum_i \sum_j^{K \times K} I_{ij}^{\text{obj}} \sum_{c \in \text{class}} \left[-p_i(c) * \log(\hat{p}_i(c)) - (1 - p_i(c)) * \log(1 - \hat{p}_i(c)) \right] \quad (9)$$

式(9)中, $p_i(c)$ 代表第 i 个网格中边界框覆盖的目标属于类别 c 的概率.

1.2 人体行为特征提取

1.2.1 3D视觉分析

传统的人体行为特征提取技术只能从RGB视频

中获取二维特征, 致使特征维度信息缺乏完整性^[7-8]. 因此, 本次研究在二维信息的基础上加以拓展, 在三维空间中提取人体行为特征.

选取一个采用RealSense相机所拍摄的RGB图像, 将RGB图像转换为深度图并对其实施伪色彩处理, 成像效果如图2所示.



(a) RGB图 (b) 深度图

图2 带有深度信息的RGB图像

RealSense相机所提供的各个像素值均含有该点与相机之间的平面距离信息, 相机坐标系点 (X, Y, Z) 与图像坐标系点 (x, y) 之间的转换关系如下:

$$\begin{cases} X = \frac{z(x - C_x)}{f} \\ Y = \frac{z(y - C_y)}{f} \\ Z = d \end{cases} \quad (10)$$

式(10)中, f 代表相机焦距记为, d 代表像素深度值记为, (c_x, c_y) 代表图像主点.

本次研究基于深度图像与RGB图像之间的互补性, 采用常规的场分割算法来标识出RGB图像的地面区域, 并将该区域统一标记为绿色, 标识效果如图3所示.



(a) 原图 (b) 标识结果

图3 地面标识效果

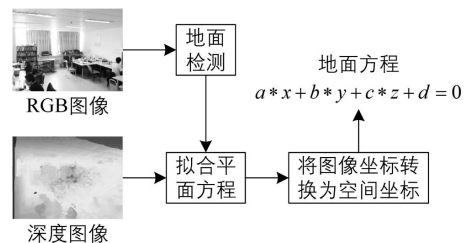


图4 地面区域方程计算流程

提取标识区域的像素坐标值与深度值, 通过式(10)将RGB二维坐标转换为三维空间坐标, 在此基础上通过常规最小二乘法拟合出平面方程参数, 获取地面区域方程,

具体流程如图4所示。

1.2.2 3D包围框

本次研究采用人体外接矩形框的纵深比来描述医院患者在跌倒状态下的外形特征,纵深比表达方式如下:

$$R_{WDH} = \frac{WD}{H} \quad (11)$$

式(11)中,WD代表人体3D包围框纵深与横向对角线的长度,在人体处于站立姿态的情况下, R_{WDH} 值将小于1;在人体处于跌倒状态的情况下, R_{WDH} 值将大于1,具体表现形式如图5所示。

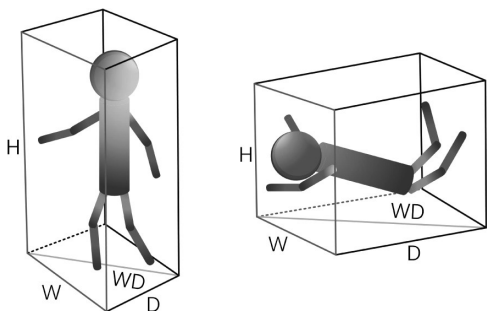


图5 人体3D包围框

1.2.3 人体距离地面的高度计算

人体的跌倒往往表现为身体在短时间内迅速贴近地面,本次研究通过人体头部与地面之间的距离来对入地之间的高度变化特征加以描述,表达方式如下:

$$h = \frac{|a * X + b * Y + c * Z + d|}{\sqrt{a^2 + b^2 + c^2}} \quad (12)$$

式(12)中,X、Y、Z代表人体头部区域3D点云的中心坐标,a、b、c、d均为地面方程系数。本次研究模拟了两个患者在跌倒状态下的头部高度变化轨迹,模拟结果如图6所示。

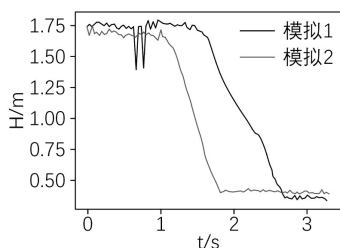


图6 模拟跌倒人体头部高度变化轨迹

根据图6可知,人体在跌倒状况下头部相对于地面的高度会急速下降,若单纯通过阈值法进行判断则难以对躺下和跌倒加以准确区分,因此需要将高度与时间两项指标结合起来建立时序运动特征,使判断方法更加具有鲁棒性。

1.3 跌倒行为判别

本次研究通过深度森林算法对人体的跌倒行为进行判别,该算法是一种用于分类问题的决策树集成算法,主要由级联森林和多粒度扫描所构成,其中多粒度描述负责获取目标特征的结构特点和上下文特点,可以借由滑动窗口机制实现针对原始输入特征的多次滑动采样^[11];级联森林是一种基于深度卷积神经网络的深度模型,能够对特征类别加以预测,该模块的输入为通过多粒度扫描所取得的特征^[12],由模块中所内置的随机森林分类器对所输入的特征进行编码处理并输出类向量,将经过多粒度扫描的目标特征与所输出的类向量相连并将其作为一下层的输入,在多次重复该操作的情况下可强化目标特征的表达效果并提升特征的多样性,最终基于最后一层输出的全部类向量计算各向量的平均值,该值即为模型所给出的预测结果。

在实际应用中,本次研究通过深度森林算法中的滑动窗口机制来获取人体跌倒状况下的时序特征,通过预置的窗口长度来捕捉时间特征序列。如图7所示。

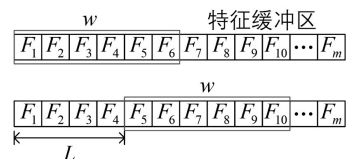


图7 滑动窗口提取特征

在图7中,L为滑动步长,w为窗口长度, F_m 为基于第m帧图像所提取的特征。在此基础上,针对目标行为发生的平均时长进行统计分析,合理设置滑动窗口的步长和尺寸,再从窗口中提取特征序列,通过级联森林模块对特征序列实施分类处理,最终实现针对跌倒行为的判别。

2 跌倒识别实验

2.1 数据样本准备

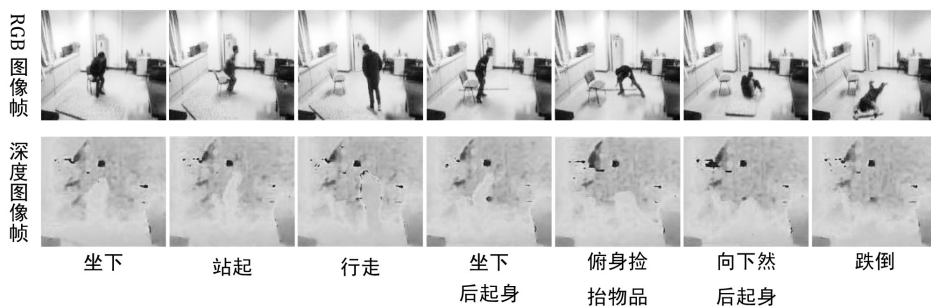


图8 七种人体行为样本示例

本次研究通过真人模拟跌倒的方式来建立实验数据集,通过RealSense D435i深度相机来捕捉目标行为图像,视频FPS为30,分辨率为640×480,分别采集行走、站起、坐下、跌倒、躺下后起身、俯身捡拾物品、坐下后起身等7种人体行为图像,共获得560组样本,其中跌倒样本中的25%用于测试,75%用于训练,样本形式如图8所示。

2.2 参数设置

经过多次跌倒模拟发现,人体从直立至跌倒所经历的时间为4s左右,考虑到帧间冗余问题及降低特征难度等方面的需求,本次研究将采样率设置为10帧/s。因此时序特征 F_m 的维度为 $4 \times 10 \times 5 = 200$ 。同时设置三个滑动窗口,窗口大小分别为150、100、50,决策树数量设置为100。为降低误检概率,分别设定阈值 $T=5s$, $H=0.25m$ 。

2.3 实验结果

将本次研究所设计的医院患者跌倒识别方案与基于二维平面特征的识别方案进行对比,所得识别结果如表1所示。

表1 不同算法下的跌倒识别结果

特征类型	误检率/%	特异性/%	灵敏度/%
3D时序特征	3.6	95.2	97.6
二维平面特征	12.5	86.9	88.1

根据对比结果可知,相较于基于二维平面特征的识别算法,基于3D视觉分析技术的医院患者跌倒识别方案在各项指标上均能够体现出显著优势。

3 结束语

本文详细介绍了3D视觉分析技术在医院患者跌倒识别中的应用思路,具体阐述了相关的样本提取方法与数据分析方法,通过对比实验的方式对基于3D视觉技术的跌

倒识别方案与常规二维识别方案进行对比,经实验研究发现,基于3D视觉技术的跌倒识别方案在灵敏度、特异性和误检率等方面均体现出显著优势,具有一定的应用价值。

参考文献:

- [1] 甘惠兴,黄鑫明.一种激光与视觉图像传感器融合的行人识别方法[J].传感器世界,2021,27(3):42-43.
- [2] 赵博.基于机器视觉图像的篮球运动训练动作捕捉方法[J].商丘师范学院学报,2021,37(6):21-22.
- [3] 李纪鑫,赫磊,任高明.基于深度学习的双目立体视觉图像特征点匹配方法[J].自动化与仪器仪表,2022(2):57-58.
- [4] 陈胜利,叶贞,罗顺权.机器视觉图像采集实验平台的开发与应用[J].机械工程与自动化,2022(1):158-159.
- [5] 赵诚,王志举,张巍,等.基于多视角车辆轮廓深度图像的割线迭代接近角计算方法[J].自动化技术与应用,2023,42(3):60-62,79.
- [6] 卢凯良.基于可见光视觉图像的路面裂缝识别深度学习方法述评[J].计算机工程与科学,2022,44(4):174-175.
- [7] 高京,董华丛.基于目标检测的铁路货车承载鞍错位故障识别[J].自动化技术与应用,2023,42(10):63-67.
- [8] 黄鼎键,杨华山,林辰耀,钟勇,王尚漳.融合超声传感与视觉图像的智能小车避障设计[J].福建工程学院学报,2021,19(6):111-112.
- [9] 杨晶,谢梅玲.基于视觉图像的多自由度工业机器人轨迹控制方法[J].制造业自动化,2021,43(11):161-162.
- [10] 李严,汪赢,韦俊.基于机器视觉的健美操跳跃动作特征提取方法研究[J].自动化技术与应用,2023,42(8):38-41.
- [11] 唐晓.基于人工智能的三维多媒体视觉图像识别研究[J].电子制作,2021(8):45-46.
- [12] 潘利利.基于体育视觉图像技术的动作识别系统构成要素及构建[J].粘接,2020,44(10):91-92.

作者简介:许庆(1991-),女,硕士,助理研究员,研究方向:医院管理。

通信作者:陈文涛(1993-),男,硕士,助理研究员,研究方向:信息管理。

(上接第74页)

参考文献:

- [1] 魏学亮,杨明顺,冯丹,等.面向安全持久性内存的元数据协同管理方法[J].计算机研究与发展,2022,59(11):2437-2450.
- [2] 吴薇薇,赵崇荣.基于元数据的图书情报信息资源整合系统[J].现代电子技术,2021,44(22):126-130.
- [3] 吴彦文,严巍,何华卿,等.融合图像元数据的用户情感分类[J].计算机工程与设计,2022,43(1):127-134.
- [4] 贺文娇,刘欣,张海瀛,等.目标情报数据质量评估[J].电讯技术,2021,61(8):932-938.
- [5] 郝如海,智勇,祁莹,等.基于熵权法的新能源高渗透率送端电网风险模糊综合评价[J].电力科学与技术学报,2021,36(1):73-78.
- [6] 柳双,杜志叶,王俊,等.基于熵权法的华中500kV高压输电线路雷击跳闸情况分析[J].现代电力,2021,38(3):332-338.
- [7] 杨寒石,郑凯文,谭纪尧.改进TOPSIS-RSR法在数控机床质量评价的应用[J].自动化技术与应用,2022,41(4):4-7.
- [8] 汪春播,王博,刘璐,等.基于元数据的开放政府数据质量自

动评估系统[J].中国管理信息化,2022,25(1):180-186.

[9] 冯维.高校创新创业课程多媒体教学资源库建设方案探讨[J].电视技术,2022,46(1):11-14.

[10] 明哲,余芸,甘杉.电力业务流转数据库中分布式数据一致性算法[J].电子设计工程,2023,31(1):120-123.

[11] 张敏,余佶,王琪琰,等.基于熵权法和灰色关联分析的藤茶理化品质评价与分级应用[J].食品与机械,2021,37(1):186-192.

[12] 朱晔,刘欣,慕小斌,等.基于层次分析和风险熵权的多站融合综合能源系统多指标综合评估[J].电测与仪表,2022,59(4):128-136,143.

[13] 谢敬东,陆池鑫,鲁思薇,等.基于序关系-熵权法的电力市场风险评估[J].中国电力,2021,54(6):71-78.

作者简介:赵东生(1984-),男,硕士,高级工程师,研究方向:金融科技和数据管理方向。