

基于三维卷积的帕金森患者拖步识别

陈晓禾^{①②} 曹旭刚^② 陈健生^{*③⑥} 胡春华^④ 马羽^⑤

^①(中国科学院苏州生物医学工程技术研究所 苏州 215163)

^②(长春理工大学电子信息工程学院 长春 130022)

^③(清华大学电子工程系 北京 100084)

^④(清华大学航天航空学院 北京 100084)

^⑤(清华大学玉泉医院 北京 100040)

^⑥(北京信息科学与技术国家研究中心 北京 100084)

摘要: 冻结步态(FoG)是一种在帕金森病(PD)中常见的异常步态,而拖步则是冻结步态的一种表现形式,也是医生用来判断患者的治疗状况的重要因素,并且拖步状态也对PD患者的日常生活有很大影响。该文提出一种通过计算机视觉来实现患者拖步状态自动识别的方法,该方法通过以3维卷积为基础的网络结构,可以从PD患者的TUG测试视频中自动识别出患者是否具有拖步症状。其思路是首先利用特征提取模块从经过预处理的视频序列中提取出时空特征,然后将得到的特征在不同空间和时间尺度上进行融合,之后将这些特征送入分类网络中得到相应的识别结果。在该工作中共收集364个正常步态样本和362个具有拖步状态的样本来构成实验数据集,在该数据集上的实验表明,该方法的平均准确率能够达到91.3%。并且其能从临床常用的TUG测试视频中自动准确地识别出患者的拖步状态,这也为远程监测帕金森病患者的治疗状态提供了助力。

关键词: 视频序列分析; 3维卷积; 异常步态识别; 拖步识别

中图分类号: TN911.73; TP181

文献标识码: A

文章编号: 1009-5896(2021)12-3467-09

DOI: [10.11999/JEIT200543](https://doi.org/10.11999/JEIT200543)

Shuffling Step Recognition Using 3D Convolution for Parkinsonian Patients

CHEN Xiaohu^{①②} CAO Xugang^② CHEN Jiansheng^{*③⑥} HU Chunhua^④ MA Yu^⑤

^①(Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou 215163, China)

^②(Department of Electronic and Information Engineering, Changchun University of Science and Technology, Changchun 130022, China)

^③(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

^④(School of Aerospace Engineering, Tsinghua University, Beijing 100084, China)

^⑤(Yuquan Hospital, Tsinghua University, Beijing 100040, China)

^⑥(Beijing National Research Center for Information Science and Technology, Beijing 100084, China)

Abstract: Freezing of Gait (FoG) is a common symptom among patients with Parkinson's Disease (PD). In this paper, a vision-based method is proposed to recognize automatically the shuffling step symptom from the Timed Up-and-Go (TUG) videos based. In this method, a feature extraction block is utilized to extract features from image sequences, then features are fused along a temporal dimension, and these features are fed into a classification layer. In this experiment, the dataset with 364 normal gait examples and 362 shuffling step examples is used. And the experiment on the collected dataset shows that the average accuracy of the best method is 91.3%. Using this method, the symptom of the shuffling step can be recognized automatically and

收稿日期: 2020-04-12; 改回日期: 2021-03-24; 网络出版: 2021-04-29

*通信作者: 陈健生 jschenthu@mail.tsinghua.edu.cn

基金项目: 国家自然科学基金(61673234)

Foundation Item: The National Natural Science Foundation of China (61673234)

efficiently from TUG videos, showing the possibility to remotely monitor the movement condition of PD patients.

Key words: Video sequence analysis; 3D convolution; Abnormal gait recognition; Shuffling step recognition

1 引言

帕金森病(Parkinson's Disease, PD)是一种常见于中老年群体的渐进性中枢神经系统退行性疾病,并且PD患者经常伴随着多种运动障碍,这些症状会严重影响患者的生活质量。在与PD患者相关的运动障碍中冻结步态(Freezing of Gait, FoG)是最常见的一种类型,其主要发生在帕金森病的中晚期,据统计60%的中晚期PD患者都表现有步态冻结的相关症状^[1-4]。在临床实践中,冻结步态的严重程度变化也能很大程度地反映患者的治疗情况,因此医生可以通过PD患者的运动状况来判断他们的病情发展^[5,6]。为了对冻结步态的运动状况进行更加详细的描述,文献^[7]根据患者的运动情况将冻结步态分为3种类型,即(1)拖步:拖曳并且小步前行;(2)原地震颤型:双腿震颤不能有效移动;(3)完全不能移动型:腿部完全不能进行有效移动。在日常生活中,原地震颤和完全不能移动两种症状的患者如果没有其他人的帮助就不能有效地活动,甚至不能单独完成医生要求的运动测试。而与之相比具有拖步症状的PD患者具有部分独立运动的能力,随之而来的便是具有拖步症状的患者更容易在日常生活中摔倒。因此,需要给予具有拖步症状的患者更多的关注来监测他们运动情况的变化,以便对治疗方案进行调整,防止患者的症状继续恶化为另外两种更严重的形式。尽管现今远程医疗的发展可以避免患者频繁到医院就诊,但是受限于宝贵的医疗资源,即使通过远程视频的方式医生也很难及时对患者的状况进行评估。所以面对这些实际的需求,本工作希望结合远程医疗的形式,提出一种能对具有拖步症状的患者进行远程自动识别的方法,以此来监测拖步患者的运动状态,帮助医生及时掌握患者的病情变化。

在以往研究异常步态识别的工作中,通过运动传感器来采集运动数据是最常用的一类方法。研究人员可以使用运动传感器采集到的数据提取出患者的多种运动参数,如行走速度、方位角等,以此来评估患者的运动状态。例如,Camps等人^[8]提出了一种通过运动传感器和深度学习技术结合的方法来识别PD患者冻结步态。首先由安装在患者腰部的惯性测量单元(Inertial Measurement Unit, IMU)采集运动信号,然后将收集到的运动数据通过8层一维卷积网络进行分类。类似的Mileti等人^[9]利用

安装在下肢的可穿戴传感器采集步态运动信号,来分析患者的运动情况。Nguyen等人^[10]则使用深度相机Kinect来获取包含运动信息的人体3维骨架,然后利用骨架信息建立人体步态模型来检测异常步态。高发荣等人^[11]通过采集下肢表面肌电信号来实现步态模式的准确识别,其结合粒子群优化(Particle Swarm Optimization, PSO)算法和支持向量机(Support Vector Machines, SVM)分类算法来对采集到的肌电信号进行分类,然后实现对行走步态的不同阶段的划分。Morris等人^[12]使用安装在人体下肢的惯性传感器借助计算机生成人体下半部分的动画模型,并统计这些人体动画模型在行走过程中的各种运动参数来对患者的冻结步态进行判断。这些基于传感器的方法通常在运动障碍的检测中有很好的性能,但是这些方法在日常中生活并没有得到广泛的应用。因为帕金森病是一种慢性运动障碍疾病,患者不方便经常去医院就诊,而且在没有专家帮助的情况下安装和校准这些传感器对普通患者来说是很困难的,另外过多的传感器也会干扰参与者的实际运动情况,因此寻找一种可以帮助医生更加简便地获得患者的运动情况的方法就非常具有意义。

随着深度学习的发展,通过计算机视觉的方法来对行为识别也吸引了很多研究。与基于传感器的方法相比,基于视觉的方法是非接触不需穿戴额外的设备,减少了患者的操作难度。Hu等人^[13]提出了一种基于视觉的帕金森病患者冻结步态识别方法,该方法首先检测患者的腿和脚的关键点,然后利用图卷积的方法来判断PD患者是否具有冻结步态。Tang等人^[14]提出了一种利用摄像机实现行走过程中关键动作检测的方法。在其工作中,步态周期被定义为从脚后跟着地开始到下次同一脚后跟再次着地的一个运动周期。其从视频帧中提取的连续轮廓差分(Continuous Silhouette Differential, CSD)特征图为脚掌离地事件检测提供了重要的特征,并以此来表示步态模式。Wolf等人^[15]利用多视点3维卷积神经网络(Multi-View 3-Dimensional Convolutional Neural Network, MV3DCNN)从步态序列中获取时空信息。其还介绍了用于处理不同服装和颜色的光流图像对识别任务的影响,其中为了解决卷积网络不能处理长序列图像的问题,将一个步态序列分割成几个短序列作为网络的输入。刘

天亮等人^[16]使用融合时空双网络流和视觉注意的方法来进行行为识别,其首先逐帧提取出视频中行为运动的光流特征,然后再使用深度学习的方法对视频图像和光流特征进行分类获得分类结果。吴培良等人^[17]提出一种视角无关的时空关联深度视频行为识别方法。其使用深度卷积神经网络将不同视角下的人体姿态映射到与视角无关的高维空间,接着对视频时间子序列进行编码,然后通过相应算法得到最终的时空特征,并根据时空特征对数据进行行为识别分类。在上述的工作中,任务目标并没有以PD患者的拖步状态研究为主要研究对象,而是对更宽泛的异常步态进行识别,所以其在拖步的识别准确率上没有达到最优,而本任务仅仅关注拖步这个现象进行研究。

3维卷积的发展给视频动作识别任务带来了许多的突破,其3维卷积核能对连续帧图像进行卷积,提取出图像序列的时间和空间特征。在3维卷积网络结构中经常使用的有C3D网络^[18]、D3D网络^[19]和P3D网络^[20]等结构,这些网络结构都在行为识别相关任务中有很好的表现,但是3维卷积网络对应的网络参数会比2维卷积更多,导致对计算机的算力有更高的要求。为了解决这个问题,本方法避免了使用层数过多的网络结构,而是结合3维卷积网络的结构和拖步状态患者的行走特点,提出了一种新的网络结构来实现拖步症状的准确识别,该网络结构大致可分为特征提取模块和特征融合模块。首先特征提取模块用于从图像序列中提取时空特征,然后再经过特征融合模块对特征进行融合判断。对于特征提取模块中的3维卷积的结构,该文选择参考最基础的C3D网络,而特征融合模块则参考Gait-

Set^[21]的网络结构以及该文中新提出网络结构对提取到的特征进行融合。该网络组合的识别准确率达到91.3%并且具有较好的鲁棒性。

因为该网络结构的输入数据是连续的“行走”图像序列,因此具有远程操作的可行性,并很容易与大多数现有的基于视觉的帕金森病患者评估方法结合起来,这也为远程医疗的发展提供了一定的助力。

2 拖步识别

在临床治疗中,医生通常使用TUG(the Timed Up-and-Go)测试^[22]来评估PD患者的基本运动功能。如图1所示,在TUG测试中患者要依次完成坐、站立、行走、转身、走回和坐回这6个涵盖了日常生活中大部分基本活动的子任务。选用TUG测试视频为基础数据,能够与现行通用的医生评估方法进行匹配,避免患者再进行额外的视频录制。

拖步识别的整体流程如图2所示,首先需要将对TUG测试视频进行自动分割获得其中的“行走”子任务,并对其进行预处理以后作为网络的输入数据。对于识别网络的结构可以分为两个模块进行描述,第1个是用于从图像序列中提取时空特征的特征提取模块,第2个则是将第1个模块得到的特征在不同的空间和时间尺度上进行融合的特征融合模块,最后使用全连接层(Fully Connected layer, FC)和sigmoid激活函数对融合后的特征进行分类识别,在本章中将分别对这几个部分进行介绍。

2.1 数据处理

本工作在清华大学玉泉医院帮助下构建了一个拖步数据集,以此作为相应的训练和测试数据,该数据集的正样本主要来自PD患者的临床TUG测试

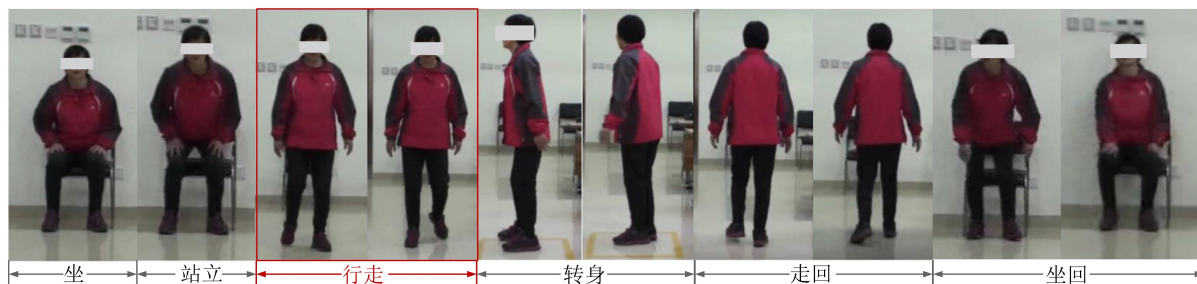


图1 TUG测试6个子任务

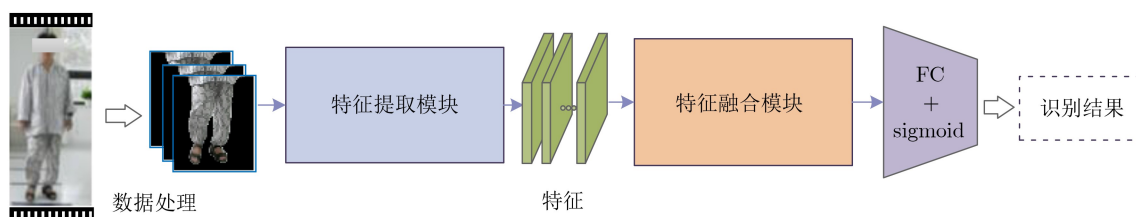


图2 整体流程方案

视频。这些TUG测试视频主要采集自24位不同年龄、性别、身高的PD患者，相关患者的信息如表1所示。每个PD患者在治疗期间需要进行深度脑刺激(Deep Brain Stimulatio, DBS)手术，而在整个治疗阶段需要进行4~6次TUG测试，并且每个患者两次TUG测试之间的间隔至少为1个月。大多数患者的TUG测试视频都能观察到拖步的症状。同时为了达到样本平衡，获得相对应的正常步态的样本，在相同的拍摄环境下录制正常人步态的TUG测试视频作为负样本。

因为我们关注的拖步与正常步态的区别主要表现在脚尖的运动差异，正常人行走过程中是先脚跟着地后脚尖着地，而对于拖步患者则表现为先脚尖着地后脚跟着地或者整个脚掌在地上拖曳前行。但是对于TUG测试6个子任务中的“站立”、“坐”和“坐回”脚部没有明显的运动，而对于“转身”和“走回”则会出现遮挡脚尖运动的情况，所以这些子任务并不能提供关于步态的有用信息。因此我们仅关注TUG测试中前向“行走”这个子任务，使用TUG视频子任务自动分割算法^[23]可以自动提取出“行走”子任务的图像序列，然后使用Mask R-CNN^[24]将“行走”片段中每一帧人体区域进行框选，同时为了进一步地消除不同视频录制时背景的干扰，使用人体语义分割方法NLGInet^[25]将人体前景与背景进行分离。因为数据集中的视频帧率为25 fps，而根据统计，正常人或具有拖步症状的PD患者在1 s内能完成一个步态周期。为了使每个输入样本中都包含一个完整的步态周期，本文选择以连续25帧图像作为一个样本，使得每个样本的持续时间为1 s。同时为了减少网络的计算量，将输入图像序列的尺寸调整为128×64。另外，因为我们的工作主要关注腿部的运动状况，而上半身的动作并不能提供太多的有用信息，所以最终仅保留图像的下1/4，得到图像的尺寸为32×64。最后，对所有样本均按照文献^[7]的标准标记为正负样本，本工作中总共收集到362个包含拖步状态的正样本和364个正常步态的负样本来组成数据集。

2.2 特征提取模块

3维卷积在视频识别中有着广泛的应用，因为其能同时从图像序列中获取空间和时间特征，本节将讨论以C3D结构为基础的3维卷积网络组成的特征提取模块。3维卷积结构中最常见的便是C3D网络，它在视频识别任务的许多数据集中得到了很好的表现。C3D的原始网络由8个3维卷积层、5个池化层和2个全连接层组成，最后使用softmax层进行分类。在文献^[18]中，实验表明C3D网络结构中卷

积核尺寸为3×3×3时有最好的表现。在本工作中参考C3D基本的网络结构得到C3D单元结构如图3所示，C3D单元仅包含1个3×3×3卷积核的3维卷积层、1个激活函数ReLU层和批归一化层(BN)，这样的简单的结构设计也能避免多层的卷积结构造成网络参数过多的问题。尽管C3D单元比大多数3维卷积网络结构简单得多，但它在实际的表现中非常优异。

在本文中，特征提取模块主要由3个阶段组成。如图4所示，特征提取模块的第1个阶段，使用1个卷积核为1×3×3的3维卷积用来提取输入图像序列的空间信息的同时保持特征 V_1 在时间维度上的独立，之后的2个阶段是使用C3D单元来提取图像序列的时空信息分别得到特征 V_2 和 V_3 。

2.3 特征融合模块

3维卷积相比于2维卷积参数会有很大程度的提升，为了控制网络的参数量，过多的卷积层并没有采用。而是结合任务的数据特性和使用场景，选择使用特征融合模块来对特征提取模块得到的各个阶段的特征进行融合处理。GaitSet在文章中提出一种利用步态序列来识别行人的方法。在GaitSet中作者提出了一个新的观点，他们把步态序列看作一组步态轮廓的集合，而行走序列中每张图片的轮廓都有其独特的外观。即使这些序列被打乱，也不难通过观察轮廓的外观将它们重新排列成为正确的顺序。因此，其假设每个轮廓都包含了其位置信息序列的顺序并不重要，而特定的姿态是区分不同步态

表1 PD患者信息统计

	平均值	范围
年龄	56.79±9.48	[37,73]
体重(kg)	63.8±10.37	[49,90]
身高(cm)	164.8±6.12	[156,178]

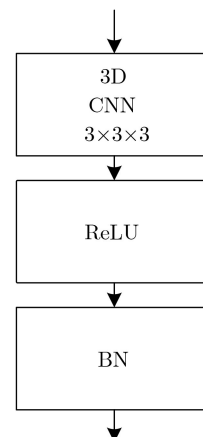


图3 C3D单元结构

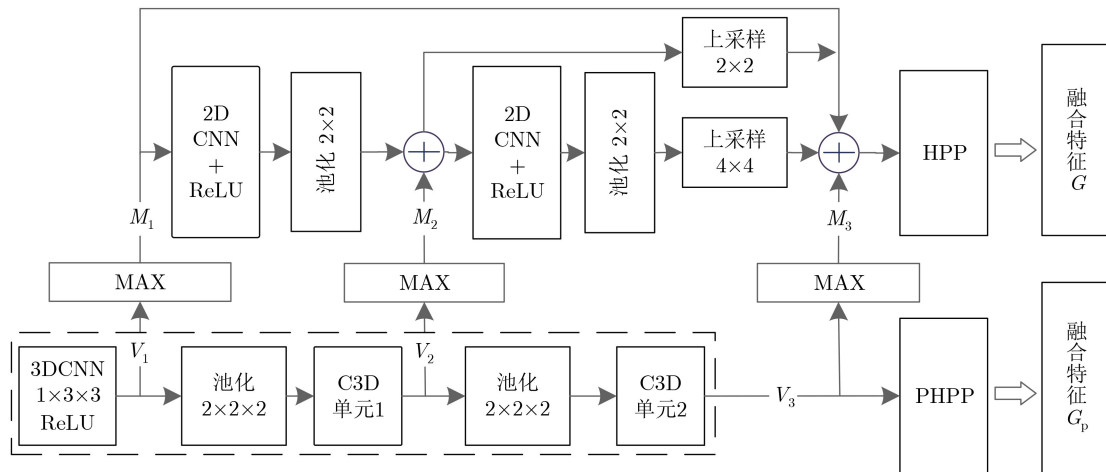


图4 识别网络结构

的主要信息。在本任务中认为区分拖步和正常步态的关键是行走图像序列中的少量帧，但是本文认为时间上的顺序是可以提供有用信息的。因此，本文参考GaitSet的网络结构并结合重新设计的网络结构来作为特征融合模块。

如图4所示，在特征融合模块中，主要使用了MAX操作、水平金字塔池化(Horizontal Pyramid Pooling, HPP)、分段水平金字塔池化(Period-wise Horizontal Pyramid Pooling, PHPP)和跨接的上采样多种操作，本文将分别对其进行介绍。通过特征提取模块得到的特征 V_1 , V_2 和 V_3 会通过MAX操作来获得其中最大响应特征并同时降低特征维度，即 $M_{(h,w,c)} = \text{MAX}[V_{(f,h,w,c)}]$ 。其中 M 代表MAX操作， $V_{(f,h,w,c)}$ 表示通过特征提取模块从连续帧图像中提取的组4维特征 V_1 , V_2 和 V_3 ， $M_{(h,w,c)}$ 表示MAX操作之后获得降维特征 M_1 , M_2 和 M_3 。之后再使用2维卷积对这些特征进一步提取，其中2维卷积的卷积核大小为 3×3 。

不同阶段的特征 M_1 , M_2 , M_3 各自包含了不同分辨率层次上的信息，为了能够保留前部分的高分辨率层次信息，在本工作中通过不同卷积之间的跨接配合不同尺寸的上采样方式将不同阶段的特征进行尺寸匹配并加和作为之后HPP^[26]的输入，其中 2×2 和 4×4 的两种采样尺寸的上采样被使用。HPP操作是通过将输入的特征进行不同尺度的池化，来充分获取输入特征中的局部和全局信息。HPP主要是使用具有不同池化尺寸的全局平均池化(Global Average Pooling, GAP)和全局最大池化(Global Max Pooling, GMP)并将不同尺寸池化后的特征进行连接。图5展示了HPP的结构，本工作中主要使用了 $(H/k, W)$ $k=1, 2, 4, 8$ 的4个不同池化尺寸对一个通道的特征进行操作，之后将各个通

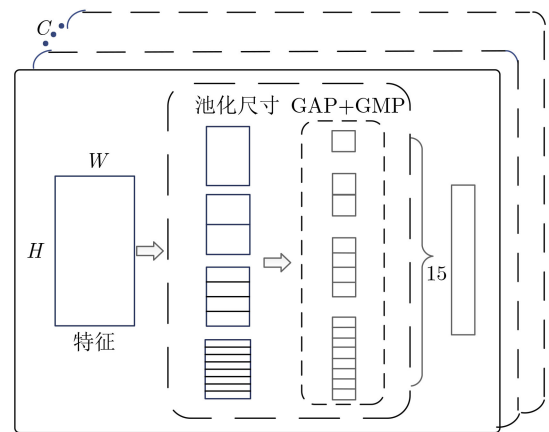


图5 HPP网络结构

道的池化特征进行连接，并使用一个全连接层对连接后的特征进行维度调整得到特征 G ，作为后续的分类特征，这种特殊的池化操作可以很好地考虑全局图片不同位置的特征信息。

而对于特征提取模块得到的特征 V_3 ，其通过多个阶段的3维卷积，其中包含输入序列的高语义信息，对网络的分类也能提供很大的帮助，因此在本文中提出一种新的结构PHPP结构来对这些特征同时进行时间和空间维度上的融合。PHPP的网络结构如图6所示，其首先将输入特征在时间维度分别划分为1, 2, 3个片段，共能够得到6个片段，然后分别对每个片段进行相同于HPP的操作，最后将每个通道的特征进行连接并使用全连接层进行维度调整得到特征 G_p 。PHPP相比于HPP保留了更多时间维度上的特征，通过在时间维度上的长序列特征和短序列划分能帮助网络获得时间维度上的全局和局部特征。

2.4 分类网络

在特征提取模块和特征融合模块之后将得到的

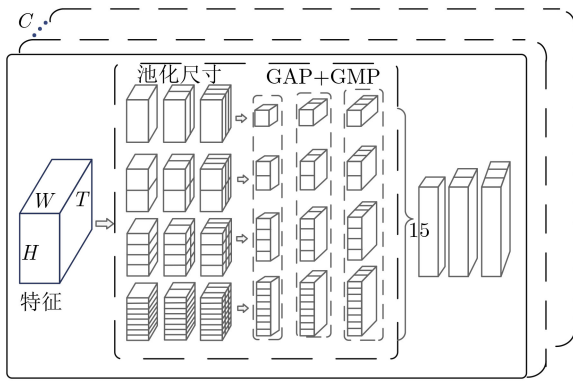


图6 PHPP网络结构

特征 G 和 G_p ，利用Flatten层将特征调整为1维，然后利用全连接层和sigmoid分类函数得到最终的结果。另外，在本文实验中，损失函数选用交叉熵函数，优化器选用Adam，并且在全连接层之前，通过添加dropout函数来提高模型的泛化能力。

3 实验与结果分析

在构建的数据集中对提出的网络结构进行测试，为了保证实验的可靠性，本工作的实验采用三折交叉验证的方式进行验证。在三折交叉验证中数据集被随机划分为3部分，并且保证同一个人的图像序列只存在一折里。在每轮交叉验证中，3个折中的一个用于测试，其余两个用于训练分类模型，最终的指标采用三折的平均值。本质上本文的任务是一个二元分类，即区分患者的各步行周期中是否包含拖步状态，在本文中使用3个指标对分类结果进行评估：精确率(precision, prec)、召回率(recall, rec)、准确率(accuracy, acc)。其中准确率定义是正确分类图像序列的百分比，而精确率、召回率计算方法分别为 $\text{prec} = \text{TP}/(\text{TP} + \text{FP})$ ， $\text{rec} = \text{TP}/(\text{TP} + \text{FN})$ ，其中TP, FP和FN分别代表真阳性(True Positive, TP)、假阳性(False Positive, TP)和假阴性(False Negative, FN)分类样本。

如表2所示，为了验证本文工作，也在本文构建的数据集上复现了几种方法。其中C3D, D3D, P3D的提出是为了视频行为识别，GaitSet则是希望通过步态识别行人的身份，JGR-GCNN则是通过视觉的方法来检测PD患者是否具有冻结步态，这也与本文的任务最为接近。其中本文方法分类准

准确率最终达到91.3%，比JGR-GCNN的准确率高出12.2%，相比其他几种方法中表现最好的准确率也提升了3.8%，并且相对应的精确率和召回率也提升很多，说明本文提出的方法不仅保持了总体的准确率，也很好平衡了该方法对正样本的敏感性和特异性。

本工作提出网络结构在识别结果中取得了很好的表现，而该文中新提出的主要结构有PHPP和跨接的上采样。表3讨论了分别消除两个操作对识别结果的影响，其中用UP来代表跨接的上采样，从表3可以看出，在消除了UP和PHPP两个结构之后，召回率会有较多的下降，但是在增加UP结构后召回率就有很大的提升，说明使用跨接的上采样操作后，在特征提取模块不同阶段得到的特征通过MAX操作之后的高分辨率特征被很好地保留，提高了网络对拖步正样本的敏感性。同时表3也展示了网络的参数量、计算量和计算时间，通过表中给出的数据可以看出，完整网络结构的计算量和计算时间都控制在非常小的范围内，证明了所提网络具有轻量化和可用性强的特点。

TUG测试视频的子任务被自动分割得到“行走”的图像序列，并被进一步处理得到多个包含完整步态周期的图像序列，这些序列以何种形式作为网络的输入也会对结果有很大的影响。如图7所示，图7(a)是正常RGB图像的形式，图7(b)采用图像轮廓的形式，图7(c)为结合图像轮廓和RGB图像得到的结果，而对于图7(d)—图7(f)分别为对应图7(a)—图7(c)仅保留腿部的结果。对于图7(b)和图7(e)采用图像轮廓的方式避免了图像背景的影响，但是很难从轮廓中观察到脚尖抬起的细微差别，图7(c)和图7(f)结合了RGB图像和轮廓图像，即减去了背景的干扰，同时保留了RGB的信息。表4

表2 不同网络比较结果(%)

网络组合	准确率	精确率	召回率
本文	91.3	89.7	92.0
C3D ^[18]	85.0	87.9	82.7
D3D ^[19]	87.5	89.4	85.1
P3D ^[20]	84.1	86.3	82.0
GaitSet ^[21]	84.9	87.1	85.2
JGR-GCNN ^[13]	79.1	87.5	76.5

表3 不同组成的网络结构试验

	准确率(%)	精确率(%)	召回率(%)	参数量(M)	FLOPs (M)	时间(ms)
无UP和PHPP结构	88.1	89.0	88.2	1.5	2.9	3.7
无PHPP结构	89.7	88.2	93.6	1.6	3.2	5.1
完整结构	91.3	89.7	92.0	1.9	3.8	7.5



图7 不同的输入图像格式

是在不同的输入序列形式下得到的实验结果，其中仅包含腿部的RGB轮廓形式有最好的结果，该输入形式既包含RGB层的信息又避免了不同背景信息的干扰还排除了上肢对识别的干扰，因此该形式输入有最好表现也符合预期设定。

4 总结

本文提出了一种使用计算机视觉来对PD患者的拖步症状进行识别的方法，在三折交叉实验中，采用一块英伟达TitanXP显卡，每一个训练轮次仅需3 min，连续训练40个轮次之后，网络的损失函数会降到最低，结果达到最优。本文为PD患者运动状态的远程监测提供了一种方案，也为PD患者

远程医疗的进一步发展提供了更多的可能性。本文提出的网络结构不仅适用于PD患者拖步的检测，对其他的异常步态识别也具有参考意义，为基于视觉的异常步态识别提出了一种新的思路。

参考文献

- [1] GBD 2016 Parkinson's Disease Collaborators. Global, regional, and national burden of Parkinson's disease, 1990–2016: A systematic analysis for the global burden of disease study 2016[J]. *The Lancet Neurology*, 2018, 17(11): 939–953. doi: [10.1016/S1474-4422\(18\)30295-3](https://doi.org/10.1016/S1474-4422(18)30295-3).
- [2] BLOEM B R, HAUSDORFF J M, VISSER J E, et al. Falls and freezing of gait in Parkinson's disease: A review of two interconnected, episodic phenomena[J]. *Movement Disorders*, 2004, 19(8): 871–884. doi: [10.1002/mds.20115](https://doi.org/10.1002/mds.20115).
- [3] DE LAU L M L and BRETHER M M B. Epidemiology of Parkinson's disease[J]. *The Lancet Neurology*, 2006, 5(6): 525–535. doi: [10.1016/S1474-4422\(06\)70471-9](https://doi.org/10.1016/S1474-4422(06)70471-9).
- [4] SVEINBJORNSDOTTIR S. The clinical symptoms of Parkinson's disease[J]. *Journal of Neurochemistry*, 2016, 139(S1): 318–324. doi: [10.1111/jnc.13691](https://doi.org/10.1111/jnc.13691).
- [5] FERRAYE M U, DEBÛ B, and POLLAK P. Deep brain stimulation effect on freezing of gait[J]. *Movement Disorders*, 2016, 31(12): 2145–2151. doi: [10.1002/mds.26600](https://doi.org/10.1002/mds.26600).

表4 多种图像输入格式实验结果比较(%)

输入数据形式	准确率	精确率	召回率
a	84.8	84.6	90.2
b	81.1	83.9	80.4
c	82.9	83.7	87.4
d	84.9	85.0	90.1
e	83.9	83.5	87.3
f	90.8	92.1	90.8

- Disorders*, 2008, 23(S2): S489–S494. doi: [10.1002/mds.21975](https://doi.org/10.1002/mds.21975).
- [6] GILADI N and HERMAN T. How do i examine Parkinsonian gait?[J]. *Movement Disorders Clinical Practice*, 2016, 3(4): 427. doi: [10.1002/mdc3.12347](https://doi.org/10.1002/mdc3.12347).
- [7] SCHAAFSMA J D, BALASH Y, GUREVICH T, et al. Characterization of freezing of gait subtypes and the response of each to levodopa in Parkinson’s disease[J]. *European Journal of Neurology*, 2003, 10(4): 391–398. doi: [10.1046/j.1468-1331.2003.00611.x](https://doi.org/10.1046/j.1468-1331.2003.00611.x).
- [8] CAMPS J, SAMÀ A, MARTÍN M, et al. Deep learning for detecting freezing of gait episodes in Parkinson’s disease based on accelerometers[C]. The 14th International Work-Conference on Artificial Neural Networks, Cadiz, Spain, 2017: 344–355. doi: [10.1007/978-3-319-59147-6_30](https://doi.org/10.1007/978-3-319-59147-6_30).
- [9] MILETI I, GERMANOTTA M, ALCARO S, et al. Gait partitioning methods in Parkinson’s disease patients with motor fluctuations: A comparative analysis[C]. IEEE International Symposium on Medical Measurements and Applications, Rochester, USA, 2017: 402–407. doi: [10.1109/MeMeA.2017.7985910](https://doi.org/10.1109/MeMeA.2017.7985910).
- [10] NGUYEN T N, HUYNH H H, and MEUNIER J. Skeleton-based abnormal gait detection[J]. *Sensors*, 2016, 16(11): 1792. doi: [10.3390/s16111792](https://doi.org/10.3390/s16111792).
- [11] 高发荣, 王佳佳, 席旭刚, 等. 基于粒子群优化-支持向量机方法的下肢肌电信号步态识别[J]. 电子与信息学报, 2015, 37(5): 1154–1159. doi: [10.11999/JEIT141083](https://doi.org/10.11999/JEIT141083).
- GAO Farong, WANG Jiajia, XI Xugang, et al. Gait recognition for lower extremity electromyographic signals based on PSO-SVM method[J]. *Journal of Electronics & Information Technology*, 2015, 37(5): 1154–1159. doi: [10.11999/JEIT141083](https://doi.org/10.11999/JEIT141083).
- [12] MORRIS T R, CHO C, DILDA V, et al. Clinical assessment of freezing of gait in Parkinson’s disease from computer-generated animation[J]. *Gait & Posture*, 2013, 38(2): 326–329. doi: [10.1016/j.gaitpost.2012.12.011](https://doi.org/10.1016/j.gaitpost.2012.12.011).
- [13] HU Kun, WANG Zhiyong, MEI Shaohui, et al. Vision-based freezing of gait detection with anatomic directed graph representation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2020, 24(4): 1215–1225. doi: [10.1109/JBHI.2019.2923209](https://doi.org/10.1109/JBHI.2019.2923209).
- [14] TANG Yunqi, LI Zhuorong, TIAN Huawei, et al. Detecting toe-off events utilizing a vision-based method[J]. *Entropy*, 2019, 21(4): 329. doi: [10.3390/e21040329](https://doi.org/10.3390/e21040329).
- [15] WOLF T, BABAEI M, and RIGOLL G. Multi-view gait recognition using 3D convolutional neural networks[C]. IEEE International Conference on Image Processing, Phoenix, USA, 2016: 4165–4169. doi: [10.1109/ICIP.2016.7533144](https://doi.org/10.1109/ICIP.2016.7533144).
- [16] 刘天亮, 谯庆伟, 万俊伟, 等. 融合空间-时间双网络流和视觉注意的人体行为识别[J]. 电子与信息学报, 2018, 40(10): 2395–2401. doi: [10.11999/JEIT171116](https://doi.org/10.11999/JEIT171116).
- LIU Tianliang, QIAO Qingwei, WAN Junwei, et al. Human action recognition via spatio-temporal dual network flow and visual attention fusion[J]. *Journal of Electronics & Information Technology*, 2018, 40(10): 2395–2401. doi: [10.11999/JEIT171116](https://doi.org/10.11999/JEIT171116).
- [17] 吴培良, 杨霄, 毛秉毅, 等. 一种视角无关的时空关联深度视频行为识别方法[J]. 电子与信息学报, 2019, 41(4): 904–910. doi: [10.11999/JEIT180477](https://doi.org/10.11999/JEIT180477).
- WU Peiliang, YANG Xiao, MAO Bingyi, et al. A perspective-independent method for behavior recognition in depth video via temporal-spatial correlating[J]. *Journal of Electronics & Information Technology*, 2019, 41(4): 904–910. doi: [10.11999/JEIT180477](https://doi.org/10.11999/JEIT180477).
- [18] TRAN D, BOURDEV L, FERGUS R, et al. Learning spatiotemporal features with 3D convolutional networks[C]. IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 4489–4497. doi: [10.1109/ICCV.2015.510](https://doi.org/10.1109/ICCV.2015.510).
- [19] LIU Jiawei, ZHA Zhengjun, CHEN Xuejin, et al. Dense 3D-convolutional neural network for person re-identification in videos[J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2019, 15(1S): 8. doi: [10.1145/3231741](https://doi.org/10.1145/3231741).
- [20] QIU Zhaofan, YAO Ting, and MEI Tao. Learning spatio-temporal representation with pseudo-3D residual networks[C]. IEEE International Conference on Computer Vision, Venice, Italy, 2017: 5534–5542. doi: [10.1109/ICCV.2017.590](https://doi.org/10.1109/ICCV.2017.590).
- [21] CHAO Hanqing, HE Yiwei, ZHANG Jumping, et al. Gaitset: Regarding gait as a set for cross-view gait recognition[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33(1): 8126–8133. doi: [10.1609/aaai.v33i01.33018126](https://doi.org/10.1609/aaai.v33i01.33018126).
- [22] PODSIADLO D and RICHARDSON S. The timed “up & go” : A test of basic functional mobility for frail elderly persons[J]. *Journal of the American Geriatrics Society*, 1991, 39(2): 142–148. doi: [10.1111/j.1532-5415.1991.tb01616.x](https://doi.org/10.1111/j.1532-5415.1991.tb01616.x).
- [23] LI Tianpeng, CHEN Jiansheng, HU Chunhua, et al. Automatic timed up-and-go sub-task segmentation for Parkinson’s disease patients using video-based activity classification[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018, 26(11): 2189–2199. doi: [10.1109/TNSRE.2018.2875738](https://doi.org/10.1109/TNSRE.2018.2875738).
- [24] HE Kaiming, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]. IEEE International Conference on Computer

- Vision, Venice, Italy, 2017: 2980–2988. doi: [10.1109/ICCV.2017.322](https://doi.org/10.1109/ICCV.2017.322).
- [25] LI Tianpeng, WAN Weitao, HUANG Yiqing, *et al.* Improving human parsing by extracting global information using the non-local operation[C]. IEEE International Conference on Image Processing, Taipei, China, 2019: 2961–2965. doi: [10.1109/ICIP.2019.8804412](https://doi.org/10.1109/ICIP.2019.8804412).
- [26] FU Yang, WEI Yunchao, ZHOU Yuqian, *et al.* Horizontal pyramid matching for person re-identification[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33(1): 8295–8302. doi: [10.1609/aaai.v33i01.33018295](https://doi.org/10.1609/aaai.v33i01.33018295).
- 陈晓禾: 男, 1976年生, 研究员, 研究方向为信号处理、人工智能.
- 曹旭刚: 男, 1995年生, 硕士生, 研究方向为深度学习与图像处理.
- 陈健生: 男, 1977年生, 副教授, 研究方向为计算机视觉与机器学习.
- 胡春华: 男, 1976年生, 副研究员, 研究方向为有源植入医疗器械.
- 马羽: 女, 1977年生, 主任医师, 教授, 研究方向为功能神经外科.

责任编辑: 马秀强