

基于混合注意力的轻量级偏瘫步态评估系统^①



余炳威, 赖舒婷, 詹润源, 郑坤升, 周成菊, 冯达钦, 潘家辉

(华南师范大学 软件学院, 佛山 528225)

通信作者: 冯达钦, E-mail: qinvsec@gmail.com

摘要: 脑卒中患者通常会出现偏瘫步态, 而视觉式步态分析可以用于检测这些变化. 然而, 当前公开的病理步态数据集规模较小、缺乏对偏瘫严重程度的详细分级, 并且传统的视觉式深度学习步态分析方法通常需要较高计算量和较大参数量, 不适用于小规模病理步态数据集. 本文设计了一款轻量级偏瘫步态评估系统. 系统使用一种轻量级卷积神经网络 (convolutional neural networks, CNN) 来评估偏瘫步态表现. 通过线性拼接不同尺度的分组卷积, 低成本地获得高效率特征. 系统引入多维度混合的轻量级注意力模块来帮助 CNN 关注空间和通道维度上的显著特征, 从而更好地平衡系统有效性与模型参数量. 此外, 本文还构建了一个专门用于步态识别的偏瘫模拟步态数据集, 为模型训练和测试提供数据支撑. 实验结果表明, 系统的神经网络仅使用 VGG-19 1/53 的参数量, 将步态识别准确率提高至 96.91%, 高于预训练后的 VGG-19, 与其他轻量化 SOTA 方法相比同样具有精度优势. 系统的开发成本低, 可部署于移动设备, 并支持实时检测, 为家庭式病理步态分析提供了一种可行的方案.

关键词: 卒中康复; 步态分析; 偏瘫数据集; 轻量级神经网络

引用格式: 余炳威, 赖舒婷, 詹润源, 郑坤升, 周成菊, 冯达钦, 潘家辉. 基于混合注意力的轻量级偏瘫步态评估系统. 计算机系统应用, 2023, 32(10):96-105. <http://www.c-s-a.org.cn/1003-3254/9260.html>

Hybrid-attention-based Lightweight Hemiplegic Gait Assessment System

YU Bing-Wei, LAI Shu-Ting, ZHAN Run-Yuan, ZHENG Kun-Sheng, ZHOU Cheng-Ju, FENG Da-Qin, PAN Jia-Hui

(School of Software, South China Normal University, Foshan 528225, China)

Abstract: Stroke patients often exhibit hemiparetic gait, and visual gait analysis can be applied to detect such changes. However, publicly available pathological gait datasets are small in scale and lack detailed grading of hemiplegia severity. Furthermore, state-of-the-art deep learning algorithms for gait analysis usually have a high need for parameter size and computational complexity, leading to low performance on small-scale pathological gait datasets. To address these challenges, this study designs a lightweight hemiplegic gait recognition system. The system utilizes an attention-based lightweight convolutional neural network (CNN) to assess hemiplegic gait performance. By linear splicing grouped convolution at different scales, high-efficiency features can be obtained at a low cost. Additionally, a multidimensional hybrid lightweight attention module is introduced to assist CNN in focusing on distinctive features in both spatial and channel dimensions, achieving a good balance between system effectiveness and lightweight design. Moreover, a hemiplegic simulation gait dataset is constructed, specifically for hemiplegic gait recognition to support model training and testing. The results demonstrate that the proposed network that uses only 1/53 parameters of VGG-19 improves the accuracy of gait recognition to 96.91%, which is higher than that of pre-train VGG-19. Compared with other lightweight SOTA methods, it also has the advantage of accuracy. The system has low development costs and can be deployed on mobile devices. It supports real-time detection, providing a feasible solution for home-based pathological gait analysis.

Key words: stroke rehabilitation; gait analysis; hemiplegia dataset; lightweight neural network

^① 基金项目: 国家自然科学基金 (62076103); 科技创新 2030 项目 (2022ZD0208900)

收稿时间: 2023-03-24; 修改时间: 2023-04-27; 采用时间: 2023-05-06; csa 在线出版时间: 2023-07-21

CNKI 网络首发时间: 2023-07-21

1 引言

脑卒中是全球性流行病之一,据2019年全球疾病负担研究(global burden of disease study, GBD)的统计数据,全球约204个国家和地区共有1.01亿例脑卒中患者,脑卒中造成1.43亿伤残调整生命(disability adjusted life year, DALY)以及655万死亡病例^[1]。脑卒中由脑血流堵塞或出血引起,在急性期过后产生偏瘫等后遗症,严重影响患者日常生活中的自理能力。然而,脑卒中康复是一个极其漫长的过程,需要大量的医疗和财力投入,为了减轻患者及其家属的负担,一个能够自动评估脑卒中严重程度以诊断中风恢复进展的低成本系统是非常有必要的。

由脑卒中引起的偏瘫,会导致行走功能障碍(也称偏瘫步态),并伴有痉挛和步态不平衡的症状^[2]。通过观察患者步态特征,我们可以评估患者的康复进度。患者的步态信息采集一般有两种方式,分别是基于可穿戴设备的步态分析^[3]和基于视觉的步态分析^[4]。基于可穿戴设备的步态分析需要在人体的每个部位安装传感器,如加速度传感器(accelerator);而基于视觉的步态分析使用成像传感器捕捉步态数据,无需穿戴设备,减轻了患者采集过程中的体力负担,提高了患者的舒适度,便于患者居家完成诊断。为了实现一个可靠的基于视觉的步态评估系统,我们需要足够多样性的数据集和高效的分类算法。

部分病理步态数据集是公开的,但由于病人的隐私问题,大多数公开数据集收集到的步态是由健康受试者模拟的。其中,数据集DAI^[5]模拟了1种正常步态和7种不同肢体摆动幅度的异常步态。INIT、GAIT-IST和GAIT-IT^[6-8]模拟了1种正常步态和偏瘫、双瘫、神经退行性疾病及帕金森4种病理性步态。以上数据集中,受试者数量最大的^[8]是21,其他^[6,7]均不超过10,大多数现有的公开数据集受试者数量规模较小,且缺乏对偏瘫步态的详细严重程度分级。两个最相关的数据集GAIT-IST^[7]和GAIT-IT^[8]均提供了严重程度级别的样本,但没有进行严重程度分级的实验。

基于视觉的步态分析算法主要可分为基于三维CNN框架和基于二维CNN框架。三维CNN^[9,10]在捕捉视频时空特征方面的有效性已被广泛证明,但由于参数过多,往往存在过拟合和收敛过慢的问题。基于二维CNN的算法缓解了三维CNN中参数量过大的问

题,它可以根据输入模态进一步分为基于序列的方法^[11]和基于模板的方法^[7,12]。基于序列的方法将连续的人体姿态序列作为输入,利用CNN对其进行处理,从中提取步态特征并进行分类或识别。相对于基于序列的方法,基于模板的方法具有更紧凑的输入,在轻量化方面具有更大的优势。大多数基于模板的方法采用步态能量图像(gait energy image, GEI)^[12]提取步态特征,但GEI容易受着装等多种协变因素的影响,留有许多无关信息;相比之下,基于骨骼低维特征的骨骼能量图像(skeleton energy image, SEI)^[7]可以避免这些影响,有效提取关键信息。综合以上分析,基于模板并将SEI作为算法输入的方法具有更大的优势。然而,在基于模板的方法中,Loureiro等人^[7]和Verlekar等人^[12]的研究都使用VGG-19^[13]进行步态分析,算法仍然包含大量模型参数,在处理小规模病理数据集时存在过拟合的问题,从而导致分类精度低。一种解决方案是先在另一个大规模数据集上进行预训练,再使用病理数据集进行微调,然而该方案需要额外的时间进行预训练,并需要另一个相关的大规模数据集,训练成本和难度较高。

针对上述问题,本文设计了一个基于视觉自动评估偏瘫步态严重程度的轻量级系统,时延低且适用于处理小规模数据集。此外,本文构建了一个面向脑卒中康复评估的以步态图像序列形式组成的数据集。基于相关医学文献^[14-19]和对真实患者步态的观察,本文将偏瘫步态分为3个等级,这比现有的公开数据集更加详细合理,且受试数量为14人,大于目前大部分病理数据集。值得注意的是,本文还通过一个轻量级注意力模块将注意力机制引入病理步态分析,以帮助网络关注通道和空间维度中的关键特征。相较于squeeze-and-excitation(SE)模块^[20]和convolutional block attention module(CBAM)^[21]两种流行的注意力分配策略,本文提出的注意力模块在有效性和轻量化之间取得了良好的平衡,是实现病理步态分析的关键。

2 相关工作

2.1 轻量级深度学习架构

传统的步态分析深度学习算法总是存在较高的计算复杂性,往往在小规模的病理数据集表现欠佳。轻量级体系结构是一种可选的解决方案。与传统的轻量级架构,如Inception^[22]、SqueezeNet^[23]对比,基于深度可分离卷积(depthwise separable convolution)的轻量化策

略在精度上更具优势,如 EfficientNet^[24]、MobileNet^[25]等.而 GhostNet^[26]则使用与之相关的深度卷积 (depthwise convolution) 进行改进,在模型的计算成本以及分类精度上达到了更好的平衡.深度可分离卷积主要包括深度卷积 (depthwise convolution) 和点卷积 (pointwise convolution) 两个过程,可以在保持有效的特征提取能力的同时降低卷积参数.深度卷积将输入的每个通道划分为独立的组,对每个组使用一个卷积核进行卷积,与普通卷积相比,其卷积核的通道数明显减小.点卷积指对深度卷积的结果进行 1×1 卷积,以弥补深度卷积造成的通道间相同空间位置信息交互的不足. GhostNet^[26] 中的 GhostModule 在深度可分离卷积之前插入了一个普通卷积以提取输入的关键特征,再将输出与深度可分离的结果进行拼接,以提高精度,且没有显著增加计算成本.然而, GhostNet^[26] 的注意力机制只探究了通道维度上的注意,忽略了空间维度上的注意,而这是找出空间特征图中的关键区域的重要因素. Woo 等人^[21]认为,空间注意力主要由池化操作和卷积操作组成,可以对空间域特征进行最大和平均池化操作,并将两个结果进行卷积操作,以此获得注意力权重.这种方法可以在不生成大量参数的情况下提取空间注意力.本文提出了基于 GhostNet^[26] 的网络框架,并对注意力模块进行了改进.

2.2 注意力模块

注意力机制可以让网络从较少的相关信息中学会去关注部分关键的信息.根据不同的注意力领域,注意力可划分为空间、通道、时间、混合注意力等.对于图像,如在本文任务中使用的 SEI 或 GEI 则包含了空间和通道维度的注意力.一些具有代表性的通道注意力网络^[20,27,28]以及由 squeeze-and-excitation network (SENet)^[20]提出的 SE 模块已被广泛应用于轻量级网络构建,如 GhostNet. SE 模块首先通过全局平均池化对信息通道维度进行压缩,然后通过两层全连接 (FC) 层学习通道间信息,最后通过非线性激活函数得到注意力权重分配.在这个模块中,两个 FC 层具有大量模型参数,使网络在轻量化上并不理想.而 efficient channel attention (ECA) 模块^[28]提出可以使用自适应一维卷积代替 FC 层,使网络更加轻量.具有混合通道和空间的注意力模块也可以提高模型的性能,如 CBAM^[21].因此,本文将在空间维度上引入轻量级通道注意力模块,以提高性能.

3 轻量级卒中评估方法

在本节中,我们将展示部署在智能手机上的卒中步态评估系统.如图 1 所示,系统包含两个分支:基于 WiFi 的远程系统和完全部署在智能手机上的本地系统.

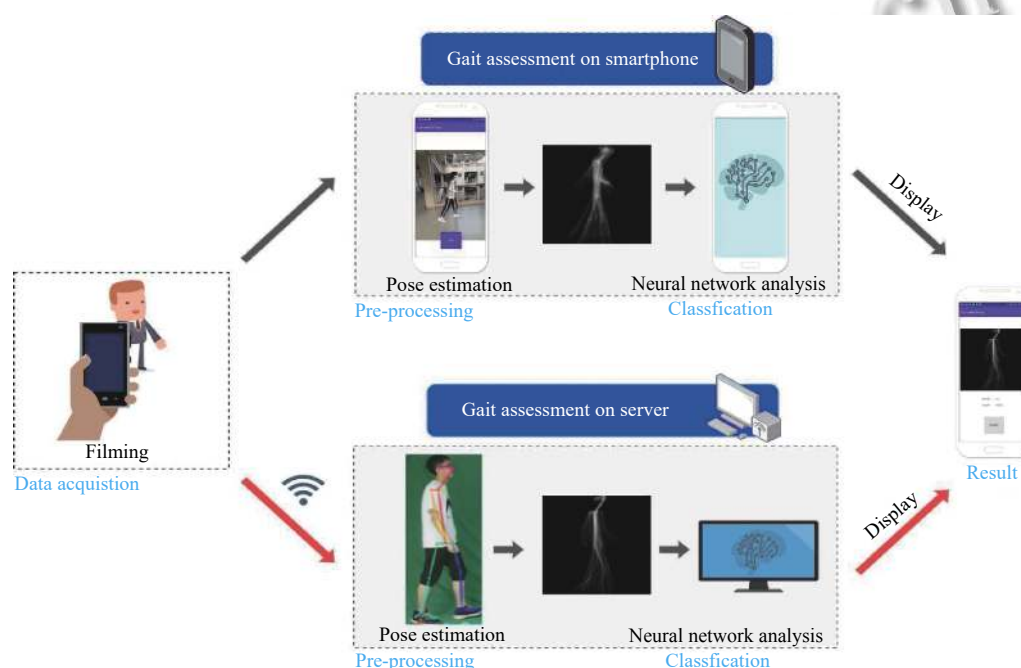


图 1 面向智能手机端的步态评估系统框架图

我们的远程系统以 APP 的形式部署, 使用 WiFi, 通过 TCP/IP 协议以及套接字技术, 将智能手机摄像头记录的步态序列传输至部署了预处理器和步态评估模型的桌面级设备, 桌面级设备再对步态序列进行分析与加工, 最后通过 WiFi 返回处理结果, APP 则通过 UI 界面显示分析结果. 相比于远程系统, 本地系统采用轻量级姿态估计方法和轻量级神经网络以及 TensorFlow Lite 框架, 无需额外的服务器和网络, 便可完全在智能手机上实现步态序列的姿态估计、SEI 生成和神经网络分类 3 个过程.

系统的两个分支均包含 3 个主要部分: 步态数据的采集, 预处理及分析. 下面, 我们首先阐述基于 WiFi 的远程系统, 并在系统实现与性能分析部分进一步讨论本地系统.

3.1 预处理

在预处理阶段, 本文首先利用人体姿态估计算法 OpenPose^[29] 从输入的彩色图像序列中提取骨骼. 然后将图像大小裁剪并统一缩放成 224×224 的分辨率, 并使人物骨骼居中. 最后, 通过步态周期检测提取骨骼序列的单个步态周期, 根据式 (1) 计算 SEI, 其中 $I_t(x, y)$ 对应于单帧图像, N 为每个步态周期的图像数. 预处理流程如图 2 所示.

$$g(x, y) = \frac{1}{N} \sum_{t=1}^N I_t(x, y) \quad (1)$$

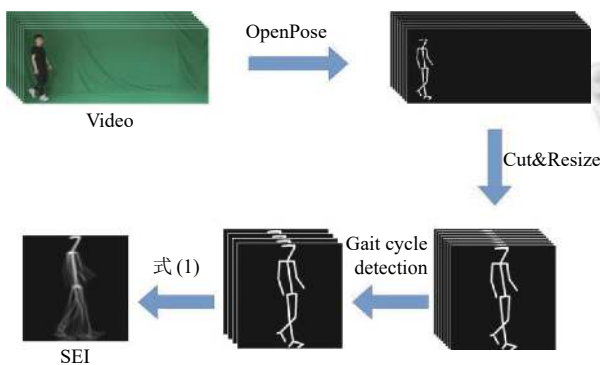


图 2 骨骼能量图像 SEI 生成过程

3.2 使用基于注意力的轻量级 CNN 进行分类

传统的基于深度学习的病理步态分析方法使用小规模数据集训练高复杂度模型, 导致高时延和过拟合. 因此, 本文设计了一种轻量级网络, 并进一步提出了一种注意力模块, 以帮助网络关注提供更多信息的空间

和通道特征, 从而进一步提高分类精度.

在多种轻量级的 CNN 算法中, GhostNet^[26] 在计算成本和准确率之间有着良好的平衡, 被选为本文的 baseline 算法. 在 baseline 的基础上, 本文提出的模型对注意力机制进行轻量化和维度混合的改进, 即改进的 GhostNet. 如图 3 所示.

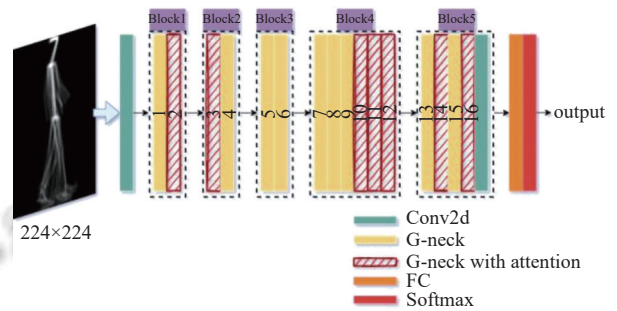


图 3 改进的 GhostNet 结构

整个网络主要包含 5 个模块, 每个模块包含数个 Ghost bottleneck (G-neck). 我们使用的 G-neck, 如图 4 所示, 由我们提出的轻量级混合注意力 (lightweight hybrid attention module) 模块和 GhostModule 组成.

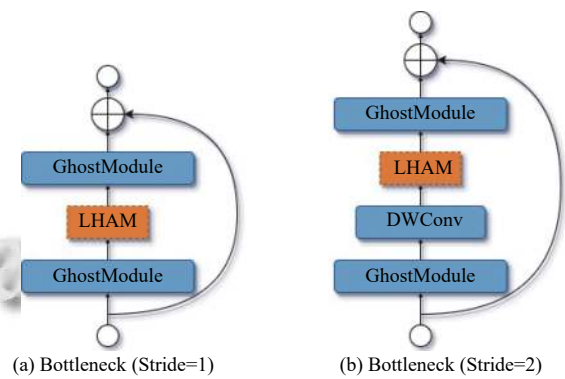


图 4 G-neck 模块

G-neck 包含两种结构, 不同之处在于其中一种包含一个额外的 DWCon (深度卷积), 以实现步幅为 2 的下采样. 本文的注意力模块, 即 LHAM, 结构如图 5 所示. 本文的模型在部分 G-neck 中放置了 LHAM 模块 (如图 3, G-neck4、5、10、11、11、12、14、16), 以关注重要的时空特征图. LHAM 优化了 GhostNet^[26] 中的 SE 模块使用两全连接层导致网络参数量过大的问题. LHAM 的具体实现是采用自适应一维 CNN 来代替网络的全连接层. 其中, 一维 CNN 的卷积核大小为 k , 它代表通道交互的覆盖范围, 与通道维度大小 C 成

正比,因此 k 可以由式(2)自适应确定,其中 odd 表示 k 的值是奇数.

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (2)$$

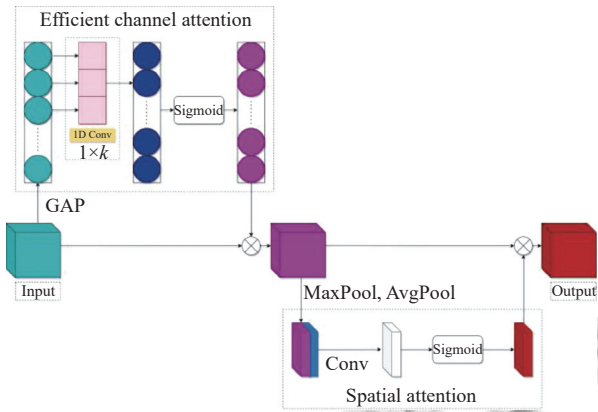


图5 本文的注意力模块

同时, GhostNet^[26]中的注意力机制仅关注在通道维度上的信息,缺少对空间维度的注意,而空间维度是在特征图中确定有效区域的关键之一.因此,对空间的注意力也包含在本文的注意力模块中.空间注意力主要由池化操作和卷积操作组成,可以对空间域特征进行最大和平均池化操作,并将两个结果进行卷积操作,以此获得注意力权值.总的来说, LHAM 采用自适应一维 CNN,大大减少了网络参数的数量.同时, LHAM 还可以在不生成大量参数的情况下有效提取空间维度的注意力,提高了网络的精度.与现有的注意力方法相比,本文提出的注意力模块实现了有效性和轻量级之间的美好平衡,进一步提升了网络的性能.

4 实验与结果

本文构建了一个偏瘫步态自构建数据集,并在该数据集以及公开的 GAIT-IST 数据集上进行了大量实验,证明了本文的系统的优越性.此外,本文还通过消融实验证明了本文提出的注意力模块的有效性.

实验按受试者分组的方式,在 GAIT-IST 数据集上进行了 10 次测试,在本文提出的数据集上进行了 7 次测试(因为数据集的受试者数量不同).

4.1 数据集构建

在本节中,本文将介绍一个用于卒中恢复的、包含以更精细等级区分的偏瘫步态自构建数据集.

具有足够受试者数量和类别多样性的数据集是步态分析算法的关键基础.表 1 列举了一些公开的病理步态数据集,并根据类别、人数和偏瘫序列的数量以及数据形式进行区分.

表 1 公共病理步态数据集

名称	类别	人数	偏瘫序列	数据模态
INIT	8	10	0	剪影
DAI2	5	5	15	剪影
GAIT-IST	9	10	90	剪影,骨骼
GAIT-IT	9	21	184	剪影,骨骼

INIT 数据集有 8 种步态类别,其中 7 种异常步态类型仅通过上肢和下肢的摆动幅度来判断,没有与偏瘫直接相关的序列. DAI2 数据集仅包含 5 种步态类别:正常、偏瘫、双瘫、神经退行性疾病和帕金森,且没有对各步态类别的严重程度有所区分. GAIT-IST 和 GAIT-IT 数据集将每种疾病类别划分为两种严重程度级别,再添加到 DAI2 以实现更精细的分类.大多数相关病理数据集均包含偏瘫步态数据,但其偏瘫序列过少,严重程度分类粗糙,不能直接应用于本文的系统.因此,本文设计了一个包含多样且合理严重程度的数据集,专门应用于卒中恢复诊断.

本文使用一个 1080p 分辨率的相机拍摄侧视图步态的彩色图像序列,并将该相机安装在一个 1.5 m 高的三脚架上,垂直于过道中心,距离为 4 m.为了方便前景分割,本文在过道旁边设置了一个绿幕.本文需要收集 1 种的正常步态和模拟 3 种不同严重程度的偏瘫步态.对于每一种步态,受试者会来回行走两次,因此每个人的一种步态可以收集 4 个序列.在收集数据前,本文参考文献 [14,30] 对偏瘫的描述,指导 14 名受试者(年龄 20-25 岁)模拟 3 种不同严重程度的偏瘫步态.整个收集过程在两天内完成,最后本文为数据集收集了 224 个偏瘫序列,且数据集具有剪影和骨骼两种数据模态.

本文的数据集根据卒中后的痉挛性程度和与痉挛性^[15-18]相关的异常协同作用模式,参照 Brunstrom^[14]的卒中康复阶段,定义了偏瘫步态的严重程度等级.异常协同作用是脑卒中后患者的一种运动障碍,表现为丧失对指定肌肉群的自我控制,以至于关节耦合运动^[31].对于下肢,异常协同作用分为伸肌协同作用(髋关节内旋、内收和伸展;膝关节伸展;以及踝关节伸展和内翻)和屈肌协同作用(外旋、外展和髋关节屈曲;膝关

节屈曲;以及踝关节屈曲和外翻)^[19].3种不同程度的模拟偏瘫步态与受异常协同作用影响的下肢面积之间的关系如图6所示.

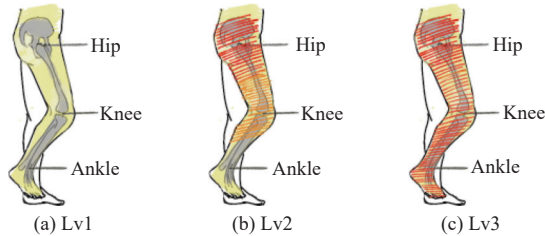


图6 下肢异常协同作用分布与受影响区域用红色斜线表示

上肢的模仿原则与下肢相似,详情请参考文献[32].每个严重程度下全身异常协同的特征描述如下.

Lv3:严重的痉挛行为,上下肢每一部分的协同作用都存在异常.在此阶段,将会模拟所有关节的异常协同作用.步态异常表现包括严重的画圈步态、髌关节伸展期间膝关节屈曲的受限以及由在下肢着地时踝关节内翻和跖屈及为弥补运动缺陷的身体侧倾造成的步态周期的双支撑相缩短.以上对应于Brunnstrom第3阶段^[14].

Lv2:痉挛开始消退,异常协同作用不明显.在此阶段,仅模拟了异常协同作用下髌关节和膝关节的高耦合运动,而踝关节的异常协同作用并没有被刻意模拟.步态异常表现包括轻度旋回以及髌关节伸展时的膝关节屈曲受限,但未达到Lv3的程度.以上对应于Brunnstrom第4到5阶段^[14].

Lv1:无明显痉挛,无异常协同作用.在此阶段,患侧上肢、下肢肢体运动缓慢,摆动幅度较小,无明显的异常协同作用.以上对应于Brunnstrom第6阶段^[14].处于这一阶段的患者可以通过适当的康复治疗来恢复正常.

4.2 SOTA分类对比

SOTA分类系统使用了微调后的VGG-19模型.该模型在ImageNet^[33]进行了预训练,并冻结第2块之前的层.本文将微调后的VGG-19模型、以及其他轻量级SOTA,如MobileNetV3、GhostNet在GAIT-IST和自构建数据集上进行了测试,并与我们的模型进行对比,实验结果如表2所示.

由表2可知,本文提出的方法在两个数据集上都优于其他SOTA方法.在GAIT-IST数据集上,本文的方法的准确性比微调后的VGG-19高0.7%,在自构建数据集上高1.56%.与微调后的VGG-19模型相比,

本文的模型参数数量减少了约53倍,每秒浮点运算次数(Flops)减少了约111倍.而且在实验之前,本文的模型不需要在大规模数据集上对本文的方法进行预训练,具有时效性.对比结果表明,该方法在分类精度和计算效率方面都有较好的性能,因此该算法更适合于小规模病理步态分析和实际实现.与其他轻量化SOTA方法的比较中,本文也使用最低参数取得了最高的准确率.

表2 步态数据集的准确性和参数量

数据集	方法	参数量 (百万)	每秒浮点运 算次数(亿)	准确率(%)
GAIT-IST	Fine-tuning VGG-19	139	196	97.40
	MobileNetV3	4.1	2.19	95.78
	GhostNet	4.2	1.77	96.24
	Our model	2.6	105	98.10
Our dataset	Fine-tuning VGG-19	139	196	95.35
	MobileNetV3	4.1	2.19	95.18
	GhostNet	4.2	1.77	95.83
	Our model	2.6	105	96.91

4.3 交叉数据集实验

对于同一数据集的实验可能会出现过度拟合的现象.为了验证本文模型的鲁棒性,本文使用所提出的模型和微调的VGG-19进行了跨数据集的偏瘫步态识别实验.

在交叉数据集实验中,本文将IST数据集作为训练集,自构建的数据集作为测试集.IST数据集共有5种步态类别(正常、神经退行性疾病、帕金森、偏瘫、双瘫),而本文的数据集包括4种步态(正常、Lv1、Lv2和Lv3).对于样本标签,本文将自构建数据集的正常步态对应于GAIT-IST数据集的正常步态,而Lv1、Lv2和Lv3步态对应于GAIT-IST数据集的偏瘫步态.如表3所示,本文的模型准确度比微调的VGG-19高出5.63%.实验结果表明,本文的模型比微调的VGG-19更具鲁棒性.

表3 在GAIT-IST数据集的准确性和参数量

方法	参数量(百万)	准确率(%)
Fine-tuning VGG-19	139	81.69
Our model	2.6	87.32

4.4 注意力模块效果的消融研究

为了验证本文提出的注意力模块的有效性,本文进行了一些消融实验,结果如表4所示.实验使用两个

流行的 SOTA 注意模块, SE 模块和 CBAM 与 LHAM 进行比较.

表 4 注意力消融实验的准确性和参数量

数据集	注意力模块	参数量(百万)	准确率(%)
GAIT-IST	SE module	4.1	97.10
	CBAM	4.1	98.13
	Our module	2.6	98.10
Our dataset	SE module	4.1	95.83
	CBAM	4.1	96.35
	Our module	2.6	96.91

实验结果显示, LHAM 的准确率要比 SE 模块高 1%, 与 GAIT-IST 数据集上的 CBAM 相当, 在本文的数据集上也分别比 SE 模块和 CBAM 高 1.08% 和 0.56%. 更重要的是, LHAM 的模块参数量要比 SE 模块和 CBAM 少 150 万. 实验结果表明, LHAM 在分类表现上要优于 SE 模块, 而在计算量上要优于 CBAM.

5 系统实现与性能分析

在本节中, 本文将深入探讨本地步态分析系统的性能, 自构建数据集的特定分类表现以及轻量级模型架构和注意力模块的影响.

5.1 本地步态评估系统性能

除了远程步态评估系统外, 本文还设计了一个无需网络或依赖其他设备, 在智能手机上便可执行的步态评估系统. 为了实现在智能手机上部署步态评估系统, 本文将神经网络从 h5 模型转换为更轻量的 tflite 模型. 而本地步态评估系统采用 MoveNet^[34] 进行姿态估计. MoveNet 的关键点要少于 OpenPose, 所提取的信息相应较少. 但是, MoveNet 是一种基于轻量级网络 MobileNet 的姿态估计算法. 它使用深度可分离卷积作为基本卷积单元, 大大减少了操作延迟和参数数量, 具有较低的计算成本和更快的速度, 以至于可以在移动设备上部署.

为了证明 MoveNet 的可靠性, 本文使用 MoveNet Lightning 处理的数据集图像输入到本文提出的模型中, 并将其与使用相同预处理方法的 OpenPose 进行比较. 由于 IST、IT 和其他公共数据集不提供原始图像数据, 本文选择自构建的数据集进行此实验. 表 5 中的实验结果表明, 与 OpenPose 相比, 使用 MoveNet 预处理的图像的分类精度降低了 1.64%. 但它仍然保持在很高的水平, 并且具有 6 倍的 FPS. 在姿态估计比较实验中, MoveNet 的准确度与 OpenPose 相当. 对 MoveNet

的关键点分布进一步分析, 我们发现它缺乏对人类踝关节的估计, 而这是判断步态是否属于 Lv3 的参数之一. 本文认为, MoveNet 之所以还能做到准确预测, 是因为在 Lv3 中有大量的上半身补偿运动, 而 MoveNet 第 6、7、12 和 13 个关节可以很好地捕捉到这些身体补偿运动, 弥补了缺少踝关节关键点而导致的准确度下降问题.

表 5 OpenPose 与 MoveNet 在自构建数据集的分类结果

姿势估计方法	准确率(%)	平均FPS
MoveNet	95.27	53.2
OpenPose	96.91	8.1

5.2 自构建数据集的步态分类结果分析

本文的模型旨在评估中风康复患者的偏瘫步态的严重程度. 本节本文将通过讨论每个步态严重程度的分类性能的详细信息, 以深入了解该模型. 在图 7 中, 本文通过混淆矩阵总结了自构建数据集上两种方法的每个类别的性能.

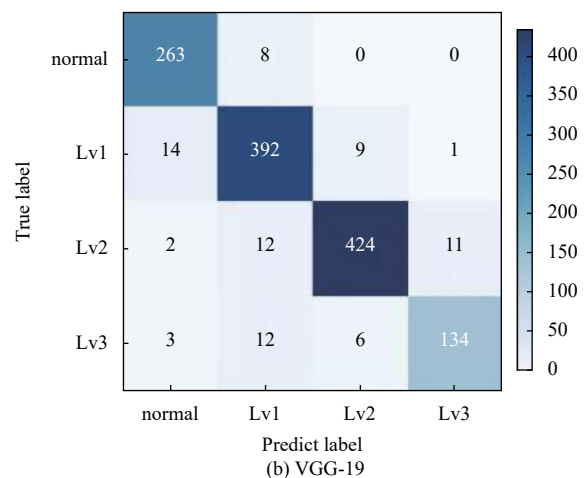
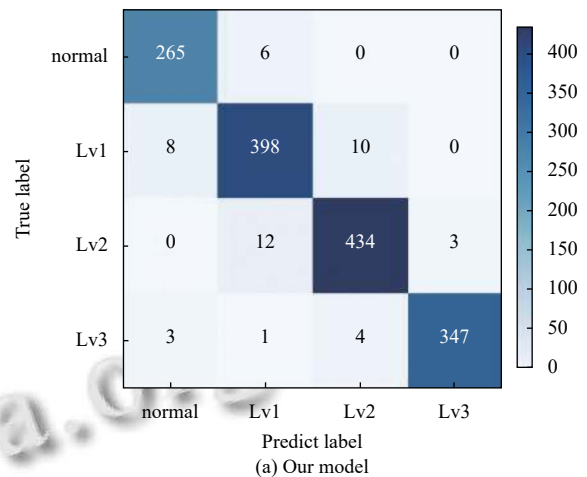


图 7 特征矩阵

如图7所示,在两种模型中,Lv1区分度均较低,容易与正常行走和Lv2混淆。相比于VGG-19,本文提出的模型更为稳固,能够提取更多的区分特征来区分不同的严重程度。与正常步态相比,Lv1步态的肢体摆动幅度和速度变化不大,容易与之混淆。仅从侧面SEI不能很好地检测到画圈步态,而画圈步态是Lv1和Lv2之间的主要区别,这可能是导致Lv1和Lv2之间出现错误分类的原因。因此,数据库补充前视图可能会对改善网络性能有所帮助。基于混淆矩阵,我们进一步观察两种模型在自构建数据集上每个标签分类的特异性和敏感性,如表6所示。两种算法的特异性相似,本文的模型相对于VGG-19改善了所有瘫痪步态类别的敏感性,这意味着我们的模型可以更为准确地检测出患者瘫痪的恢复进展,以确保对患者进行及时的医疗干预。

表6 两种模型在自构建数据集上的特异性和敏感性(%)

类别	Fine-tuning VGG-19		Our model	
	Sensitivity	Specificity	Sensitivity	Specificity
正常	97	98	97	99
Lv1	94	97	95	98
Lv2	94	98	96	98
Lv3	94	99	97	99

5.3 与SOTA步态分类表征能力的差异对比

对于相同的输入图像,图8展示了从VGG-19的第1个块中输出的特征图,该模型采用了普通的卷积架构,共包含64个通道,大小为 112×112 ,而图9展示了从本文提出的模型第1个块中输出的特征图,该模型采用了更轻量级的卷积架构,共包含24个通道,大小为 56×56 。

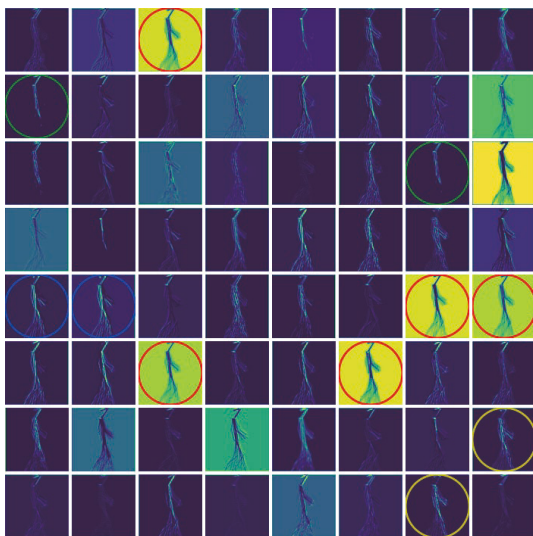


图8 VGG-19的第1个块输出的特征图

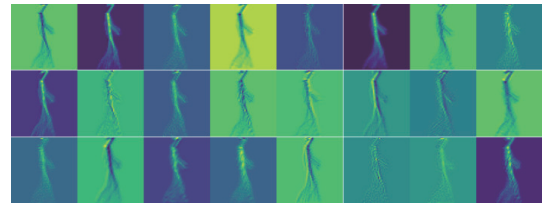


图9 本文模型的第1个块输出的特征图

特征图提取到的信息与其大小与数量有关,特征图数量越多,尺寸越大,则需要更多参数和浮点运算。通过观察图8,可以发现,图8中存在大量高度相似或几乎相同的特征图,如图8中被同一颜色圈出的特征图所示,这将导致网络中存在大量冗余信息。相比之下,本文提出模型的输出通道更少、大小更小。并且特征图的重复性进一步降低,更加精细,信噪比更高,这对于只能提供有限信息的小型数据集非常重要。

5.4 不同类别输入的空间注意力权重变化

图10展示了来自同一受试者的3种不同程度的偏瘫步态以及正常步态输入所分配的空间注意力权重矩阵热力图。

与正常步态相比,在Lv1步态中,空间注意力主要集中在下肢,与本文忽视上肢摆动的异常的医学原则相对应,这是因为上肢部分的小幅度摆动往往会与躯干重叠,不易获取有效信息。与Lv1相比,由于在Lv2中添加了画圈步态的特征,下肢的注意力权重显著增加,符合我们之前提出的医学原则。Lv3的注意力权重分布主要集中在上半身、踝部及以下部分,其在行走时更充分地利用了上半身的肌肉,这与Lv3中本文踝部异常和显著的躯干补偿的医学原则相一致。

6 结语

本文使用自构建的数据集和提出的轻量级注意力卷积神经网络部署了一种新颖的基于视觉的偏瘫步态评估系统,为在家庭式康复中对偏瘫患者进行低成本高舒适度的偏瘫严重性评估提供了一种解决方案。在公共数据集GAIT-IST和自构建数据集进行的广泛实验和详细的可视化分析表明,本文的方法具有比现有方法更低的计算成本和更高的分类准确率。

在未来的工作中,我们计划将前视角的步态数据添加到数据集中,并使用多分支轻量级神经网络融合前视图和侧视图进行分析。此外,我们的团队目前正将系统迁移到云端,以便患者家属进行远程和实时的步态观察和监测。

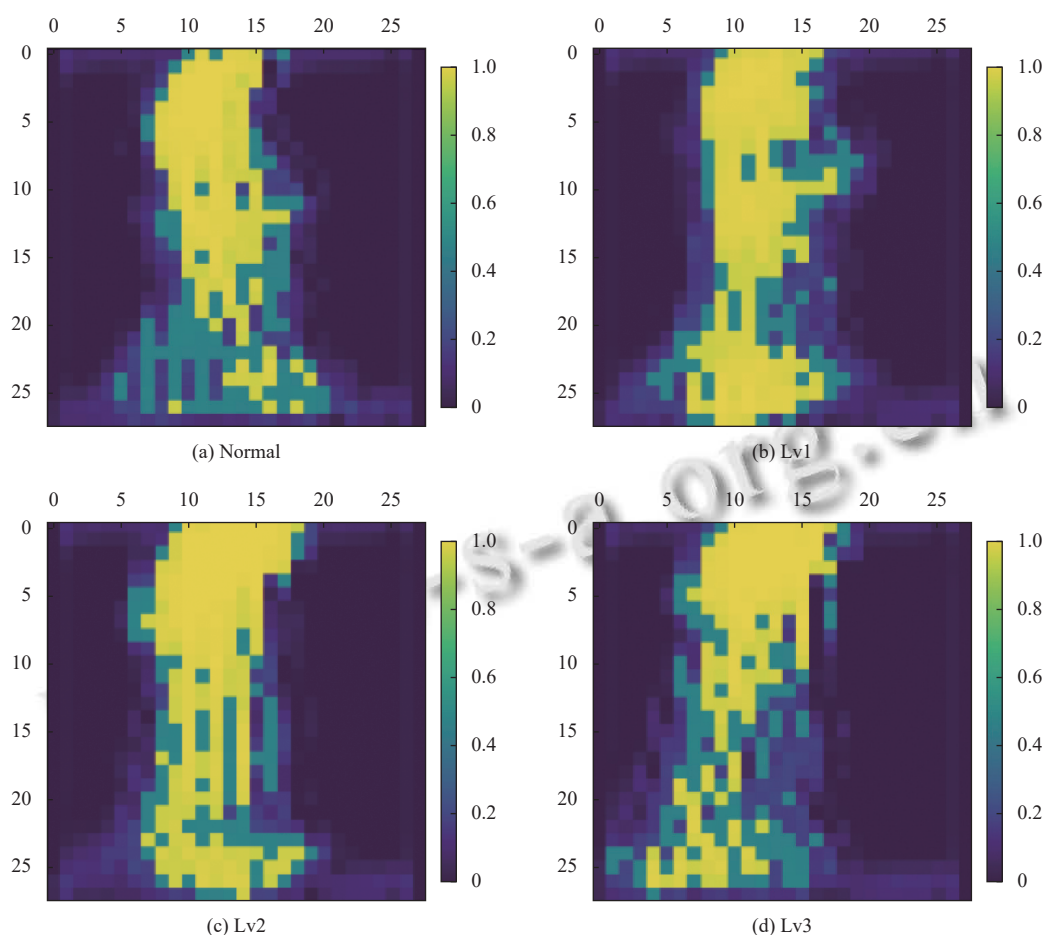


图10 不同类别空间注意力权重分布

参考文献

- 1 GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990–2019: A systematic analysis for the global burden of disease study 2019. *The Lancet Neurology*, 2021, 20(10): 795–820. [doi: [10.1016/S1474-4422\(21\)00252-0](https://doi.org/10.1016/S1474-4422(21)00252-0)]
- 2 王桂茂, 齐瑞, 严隽陶. 中风偏瘫步态的生物力学及其运动学特征分析. *中国组织工程研究与临床康复*, 2007, 11(40): 8169–8172.
- 3 Peters DM, O'Brien ES, Kamrud KE, *et al.* Utilization of wearable technology to assess gait and mobility post-stroke: A systematic review. *Journal of NeuroEngineering and Rehabilitation*, 2021, 18(1): 67. [doi: [10.1186/s12984-021-00863-x](https://doi.org/10.1186/s12984-021-00863-x)]
- 4 Stenum J, Rossi C, Roemmich RT. Two-dimensional video-based analysis of human gait using pose estimation. *PLoS Computational Biology*, 2021, 17(4): e1008935. [doi: [10.1371/journal.pcbi.1008935](https://doi.org/10.1371/journal.pcbi.1008935)]
- 5 Nieto-Hidalgo M, Ferrández-Pastor FJ, Valdivieso-Sarabia RJ, *et al.* Vision based extraction of dynamic gait features focused on feet movement using RGB camera. *Proceedings of the 1st International Conference on Ambient Intelligence for Health*. Puerto Varas: Springer, 2015. 155–166.
- 6 Ortells J, Herrero-Ezquerro MT, Mollineda RA. Vision-based gait impairment analysis for aided diagnosis. *Medical & Biological Engineering & Computing*, 2018, 56(9): 1553–1564.
- 7 Loureiro J, Correia PL. Using a skeleton gait energy image for pathological gait classification. *Proceedings of the 15th IEEE International Conference on Automatic Face and Gesture Recognition*. Buenos Aires: IEEE, 2020. 503–507.
- 8 Albuquerque P, Machado JP, Verlekar TT, *et al.* Remote gait type classification system using markerless 2D video. *Diagnostics*, 2021, 11(10): 1824. [doi: [10.3390/diagnostics11101824](https://doi.org/10.3390/diagnostics11101824)]
- 9 Lin BB, Zhang SL, Bao F. Gait recognition with multiple-temporal-scale 3D convolutional neural network. *Proceedings of the 28th ACM International Conference on Multimedia*. Seattle: ACM, 2020. 3054–3062.
- 10 Thapar D, Jaswal G, Nigam A, *et al.* Gait metric learning siamese network exploiting dual of spatio-temporal 3D-CNN intra and LSTM based inter gait-cycle-segment features.

- Pattern Recognition Letters, 2019, 125: 646–653. [doi: [10.1016/j.patrec.2019.07.008](https://doi.org/10.1016/j.patrec.2019.07.008)]
- 11 Liao RJ, Cao CS, Garcia EB, *et al.* Pose-based temporal-spatial network (PTSN) for gait recognition with carrying and clothing variations. Proceedings of the 12th Chinese Conference on Biometric Recognition. Shenzhen: Springer, 2017. 474–483.
 - 12 Verlekar TT, Correia PL, Soares LD. Using transfer learning for classification of gait pathologies. Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine. Madrid: IEEE, 2018. 2376–2381.
 - 13 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Proceedings of the 3rd International Conference on Learning Representations. San Diego: ICLR, 2015.
 - 14 Brunnstrom S. Motor testing procedures in hemiplegia: Based on sequential recovery stages. Physical Therapy, 1966, 46(4): 357–375. [doi: [10.1093/ptj/46.4.357](https://doi.org/10.1093/ptj/46.4.357)]
 - 15 Park C, Oh-Park M, Bialek A, *et al.* Abnormal synergistic gait mitigation in acute stroke using an innovative ankle-knee-hip interlimb humanoid robot: A preliminary randomized controlled trial. Scientific Reports, 2021, 11(1): 22823. [doi: [10.1038/s41598-021-01959-z](https://doi.org/10.1038/s41598-021-01959-z)]
 - 16 Neckel N, Pelliccio M, Nichols D, *et al.* Quantification of functional weakness and abnormal synergy patterns in the lower limb of individuals with chronic stroke. Journal of NeuroEngineering and Rehabilitation, 2006, 3: 17. [doi: [10.1186/1743-0003-3-17](https://doi.org/10.1186/1743-0003-3-17)]
 - 17 Beyaert C, Vasa R, Frykberg GE. Gait post-stroke: Pathophysiology and rehabilitation strategies. Neurophysiologie Clinique/Clinical Neurophysiology, 2015, 45(4–5): 335–355.
 - 18 Riad J, Coleman S, Lundh D, *et al.* Arm posture score and arm movement during walking: A comprehensive assessment in spastic hemiplegic cerebral palsy. Gait & Posture, 2011, 33(1): 48–53.
 - 19 Brunnstrom S. Movement Therapy in Hemiplegia: A Neurophysiological Approach. New York: Harper and Row, 1970.
 - 20 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
 - 21 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 3–19.
 - 22 Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 1–9.
 - 23 Iandola FN, Han S, Moskewicz MW, *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. arXiv:1602.07360, 2016.
 - 24 Tan M, Le Q. EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning. Long Beach: PMLR, 2019. 6105–6114.
 - 25 Howard A, Sandler M, Chen B, *et al.* Searching for MobileNetV3. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2020. 1314–1324.
 - 26 Han K, Wang YH, Tian Q, *et al.* GhostNet: More features from cheap operations. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 1577–1586. [doi: [10.1109/CVPR42600.2020.00165](https://doi.org/10.1109/CVPR42600.2020.00165)]
 - 27 Qin ZQ, Zhang PY, Wu F, *et al.* FcaNet: Frequency channel attention networks. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 763–772.
 - 28 Wang QL, Wu BG, Zhu PF, *et al.* ECA-Net: Efficient channel attention for deep convolutional neural networks. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 11531–11539. [doi: [10.1109/CVPR42600.2020.01155](https://doi.org/10.1109/CVPR42600.2020.01155)]
 - 29 Cao Z, Hidalgo G, Simon T, *et al.* OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(1): 172–186. [doi: [10.1109/TPAMI.2019.2929257](https://doi.org/10.1109/TPAMI.2019.2929257)]
 - 30 Stanford School of Medicine. Gait abnormalities. <https://stanfordmedicine25.stanford.edu/the25/gait.html>. [2022-05-30].
 - 31 Sakuma K, Ohata K, Izumi K, *et al.* Relation between abnormal synergy and gait in patients after stroke. Journal of Neuroengineering and Rehabilitation, 2014, 11: 141. [doi: [10.1186/1743-0003-11-141](https://doi.org/10.1186/1743-0003-11-141)]
 - 32 McMorland AJC, Runnalls KD, Byblow WD. A neuroanatomical framework for upper limb synergies after stroke. Frontiers in Human Neuroscience, 2015, 9: 82.
 - 33 Deng J, Dong W, Socher R, *et al.* ImageNet: A large-scale hierarchical image database. Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE, 2009. 248–255.
 - 34 Google Research. <https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html>. [2022-07-18].

(校对责编: 孙君艳)