

文章编号:1671-6833(2024)04-0046-07

基于混合特征图卷积神经网络的人体行为识别方法

李志新¹, 商樊淇¹, 郇战¹, 陈瑛¹, 梁久祯²

(1. 常州大学 微电子与控制工程学院, 江苏 常州 213000; 2. 常州大学 计算机与人工智能学院, 江苏 常州 213000)

摘要: 基于可穿戴传感器的人体行为识别方法不能很好地处理时间序列数据采样点之间的结构信息, 也忽略了数据样本之间的潜在联系。针对这一问题, 提出了混合时频特征和结构特征的图卷积神经网络模型进行人体动作识别。首先, 通过小波包变换获取原始信号的时频特征, 进一步构建时空图提取信号的结构特征以挖掘采样点间的动态特性, 并在结构特征中加入距离约束, 弱化时空图中远距离邻居对中心节点的影响。其次, 考虑到结构特征提取时受时空图拓扑关系影响较大, 选择样本的时频特征构造图卷积神经网络的输入拓扑, 混合时频特征和结构特征作为网络输入特征。最后, 输入特征沿着输入拓扑结构传播, 得到最终分类结果。为了评估所提模型的性能, 在 WHARF 和 DataEgo 数据集上进行了实验验证。实验结果表明: 所提模型的 $F1$ 分数相比已有的基于卷积神经网络模型在 WHARF 和 DataEgo 上均有提升, WHARF 数据集上 $F1$ 最高提升 19.58 百分点, DataEgo 数据集上 $F1$ 最高提升 26.44 百分点, 证明所提出模型通过挖掘动态特性能够有效提高动作识别能力。

关键词: 图卷积神经网络; 可穿戴设备; 人体行为识别; 时空图; 特征提取

中图分类号: TP391 **文献标志码:** A **doi:** 10.13705/j.issn.1671-6833.2024.04.002

青少年运动损伤、人群异常检测以及老年人独居安全问题受到社会的广泛关注。传感器因其采集、存储数据量少, 不涉及隐私问题以及易于集成在智能设备中的优势, 被广泛应用于行为分析^[1]、医疗保健^[2]、智能家居^[3]等人类行为感知领域。在过去的几年里, 机器学习算法吸引了人类行为识别 (human activity recognition, HAR) 领域研究者的目光。Chen 等^[4]提出了一个卷积神经网络 (convolutional neural networks, CNNs) 模型, 对卷积核进行修改以处理智能手机的三轴加速度计信号来识别人类活动。Ignatov^[5]将 CNNs 提取的局部特征与时间序列的统计特征相结合, 保留关于原始信号的全局特征信息。Wang 等^[6]提出用分层深度长短时记忆神经网络提取时序信号的时域和频域特征完成分类任务。Su 等^[7]提出深度双向长短时记忆模型将长序列原始数据处理成双向输出向量, 然后使用 CNNs 从输出向量中提取局部特征。基于长短时记忆神经网络的方法考虑到了时间序列的上

下文的语义信息, 却忽略了同种动作样本之间的相关性。此外, 这些传统深度学习方法特征提取集中在数据值上, 缺乏对数据中隐藏的结构与关系信息的探索。

与视频或时间序列信号不同, 非欧空间的图结构包含了丰富的拓扑信息。图卷积神经网络 (graph convolutional neural networks, GCNs) 因其强大的推理能力和良好的解释性成为深度学习的焦点^[8]。目前, 很少有研究者利用 GCNs 来学习和挖掘可穿戴传感器采集的人类行为信号的隐藏拓扑关系。Ahmad 等^[9]首次提出采用时间序列图模块将原始 HAR 转换为图结构, 通过 GCNs 聚合更新样本特征实现动作识别。这种图形构造方法仅通过构建样本采样点间的时序拓扑关系来完成信息传递, 没有形成样本的幅值结构信息。Donner 等^[10]总结了 3 种将时间序列转换为图形的表示方法: 递归网络、转移网络和可视图。水平可视图 (horizontal visibility graph, HVG)^[11]作为可视图的简化, 将时间序列转

收稿日期: 2023-11-16; 修订日期: 2023-12-28

基金项目: 国家自然科学基金资助项目 (62201093); 常州市科技计划资助项目 (CJ20235026); 江苏省研究生科研与实践创新计划项目 (KYCX23_3070)

作者简介: 李志新 (1993—), 女, 河南洛阳人, 常州大学讲师, 博士, 主要从事数据挖掘研究, E-mail: lizhixin2021@cczu.edu.cn。

引用本文: 李志新, 商樊淇, 郇战, 等. 基于混合特征图卷积神经网络的人体行为识别方法 [J]. 郑州大学学报 (工学版), 2024, 45(4): 46-52. (LI Z X, SHANG F Q, HUAN Z, et al. Human activity recognition based on hybrid feature graph convolutional neural network [J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(4): 46-52.)

换为图形的同时保留了时间序列拓扑与动态特征。Li 等^[12]通过加权水平可视图的方式为故障信号构建时空图,并通过图同构网络对图形进行分类。这种分类方法考虑了采样点之间的结构信息,却忽视了样本之间的信息交流。

为了解决这些挑战,本文提出了一种混合结构特征和时频特征的 GCNs (graph convolutional neural networks with hybrid time-frequency and construction characteristics, HTC-GCN)。其中心思想是将 HVG 构建的时空图转换为节点,并且在具有相似特征的节点之间构建边,实现样本自身以及样本之间的信息传递。

本文的主要工作包括以下 3 个方面。

(1)将 HVG 构建的时空图表示为拉普拉斯矩阵并通过正交分解得到动作样本的结构特征以挖掘采样点间的动态特性。

(2)在人体动作时间序列的结构特征中加入距离约束,弱化时空图中远距离邻居对中心节点的影响。进一步提高模型的拟合能力。

(3)设计了一个混合时频特征和结构特征的

GCNs 模型,克服结构特征受时空图拓扑关系影响较大的问题,进一步提高模型的识别能力。

1 人体行为识别方法

本文工作旨在建立一个 HTC-GCN 模型(如图 1 所示)突破传统 HAR 方法中时序信号结构特征提取困难的问题,探索 GCNs 在可穿戴式 HAR 领域的有效性。HTC-GCN 核心思想是将非结构化时间序列转化为语义丰富的结构化图,不仅能提取动作信号中的时频特征,还能将信号中潜在的结构信息转变为结构特征。为此,本文首先通过小波包变换提取动作信号中的时频特征 F_d , 同时采用拉普拉斯矩阵表示 HVG 算法构造的时空图,并利用正交分解提取动作信号的结构特征 F_c 。其次,时频特征 F_d 与结构特征 F_c 结合构成图神经网络输入特征 X 。此外,考虑到动作样本的划分通过滑动窗口实现导致 HVG 构建的样本拓扑存在差异,本文利用时频特征构造 GCNs 的输入拓扑。最后,通过训练两层图卷积模块,使输入特征 X 在输入拓扑图上传播,最终得到与分类任务最相关的信息 H 。

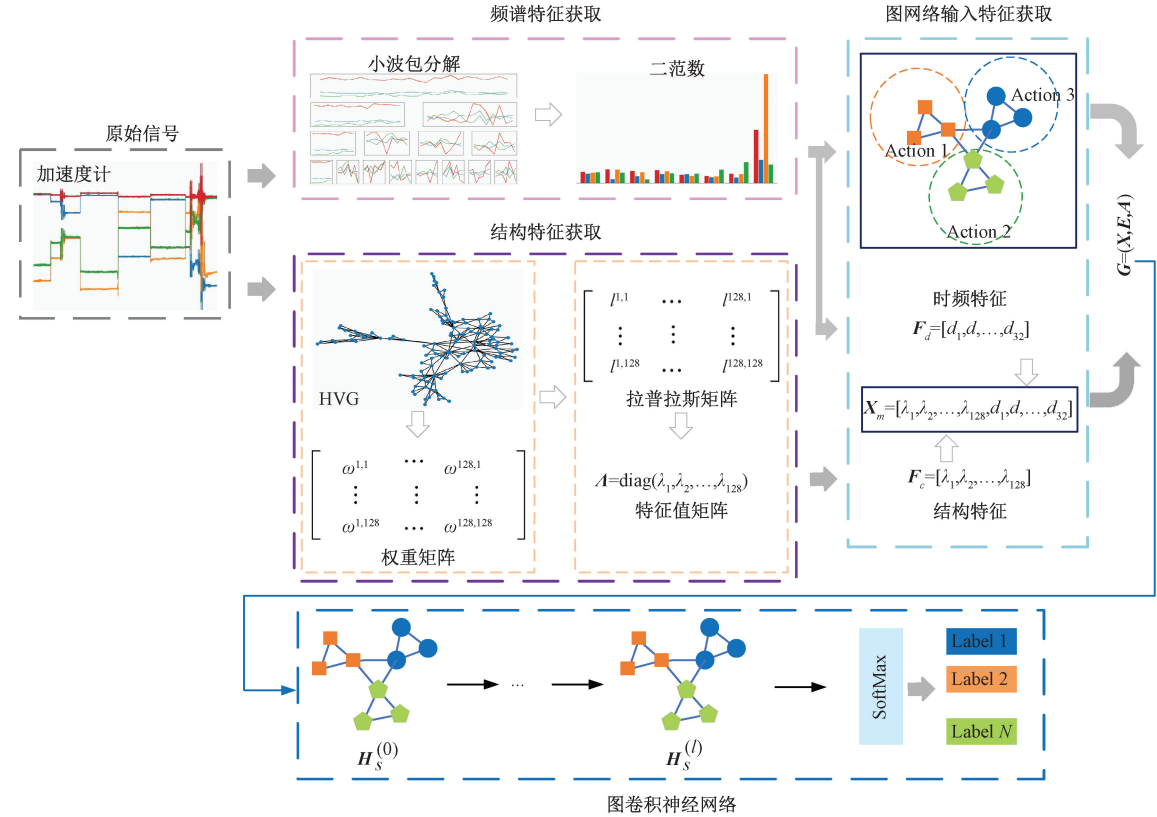


图 1 所提出的人类行为识别模型

Figure 1 The proposed HAR model

1.1 提取动作的时频特征

本文采用长度为 128、重叠率为 50% 的固定宽度滑动窗口对数据集中的加速度计信号进行处理。

使用加速度计传感器的 4 个通道 (X 轴加速度、Y 轴加速度、Z 轴加速度和三轴合成加速度), 因此将有 1 个 4×128 的输入矩阵。基于离散小波包变换

(DWPT) 对每个通道进行分解, 小波变换公式^[13-14]为

$$\begin{cases} d_i^{j,2n} = \sum_k h_{k-2l} d_k^{j-1,n}; \\ d_i^{j,2n+1} = \sum_k g_{k-2l} d_k^{j-1,n}. \end{cases} \quad (1)$$

式中: d 为小波包分解系数; h 为高通滤波器系数; g 为低通滤波器系数; j 和 n 均为小波包节点数。通过二范数计算子带长度得到时频特征 F_d 。

1.2 提取动作的结构特征

1.2.1 基于水平可视图的时空图构建

人的运动信号中包含大量的动态属性, 这些动态属性可以表示为时序信号的结构信息。本文通过 HVG 构建时空图来获取原始信号结构信息。相比于传统的特征提取方法, 时空图不仅包含了时序信号的采样点信息, 同时保留了动态属性, 进一步提高特征的识别能力。为了减弱传感器安装所造成的方向误差, 本文仅对样本的三轴合成加速度构造时空图。

(1) 图形表示。时空图表示为 $G_s = (V_s, A_s, W_s)$, 其中 V_s 为节点特征矩阵; A_s 为节点之间的边连接矩阵; W_s 为图的权重矩阵。

(2) 图形构造。HVG 算法将时间序列转换为 N 个节点的图, 节点特征为对应样本点的采样值即 $V_s = (v_s^1, v_s^2, \dots, v_s^N)$ 。仅当没有中间数据高度与 v_s^i 和 v_s^j 中间的水平连接相交时, 2 个节点之间存在无向边, 边连接矩阵 A_s 的表达如下所示:

$$a_s^{i,j} = \begin{cases} 1, & \text{节点 } v_s^i \text{ 和 } v_s^j \text{ 存在边连接;} \\ 0, & \text{其他。} \end{cases} \quad (2)$$

式中: $a_s^{i,j}$ 为 A_s 第 i 行 j 列的元素。

图 2 为 HVG 算法图解。图 2(a) 描绘了时间序列 (0.8, 0.2, 0.3, 0.6, 0.1, 0.7) 的条形图, 并画出 2 个条形柱之间未出现相交的水平线, 由此可生成如图 2(b) 所示的 HVG, 其中包含 6 个节点和 9 条无向边。由此可见, HVG 中的每个节点至少与其时间相邻的节点是有边连接的。此外, HVG 在仿射映射下保持不变, 使得不同的采样频率和归一化不会影响图形结构^[13]。

(3) 权重矩阵构造。由于传感器采集的时间序列信号中存在噪声, 导致远距离采样点之间出现干扰边缘, 这将阻碍识别性能。为了抑制干扰边缘产生的影响, 本文为时空图边连接添加如式 (3) 所示的权重约束:

$$w^{i,j} = \frac{1}{|i - j|}. \quad (3)$$

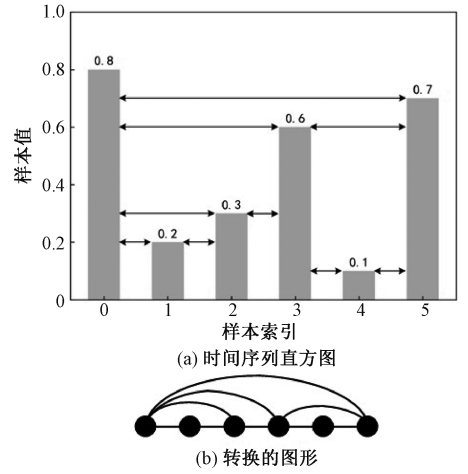


图 2 HVG 算法图解

Figure 2 Illustration of HVG algorithm

式中: i 为中心节点的时序序号; j 为邻居节点的时序序号。权重矩阵 W_s 元素由式 (4) 计算得到:

$$w_s^{i,j} = \begin{cases} \text{Dis}(v_s^i, v_s^j) w^{i,j}, & a_s^{i,j} = 1; \\ 0, & a_s^{i,j} = 0. \end{cases} \quad (4)$$

式中: $\text{Dis}(\cdot)$ 为节点 v_s^i 和节点 v_s^j 间的欧氏距离。

1.2.2 基于时空图的结构特征提取

利用图中包含隐藏结构信息的优点, 采用图论的方法进行时空图特征提取。首先依据时空图中的权重矩阵得到拉普拉斯矩阵, 然后对拉普拉斯矩阵进行正交分解得到动作的结构特征。

(1) 拉普拉斯矩阵构建。在谱图分析中, 拉普拉斯矩阵 L 的定义为

$$L = D - W_s. \quad (5)$$

式中: D 为度矩阵, 其对角线元素表示为

$$d_{i,i} = \sum_{j=1}^N \sum_{i=1}^N w_s^{i,j}. \quad (6)$$

(2) 正交分解。对实对称拉普拉斯矩阵通过式 (7) 进行正交分解:

$$L = U \Lambda U^T. \quad (7)$$

式中: $\Lambda = \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_{N-1})$ 为特征值矩阵; U 为特征向量。最后将每个时空图转变为一维特征向量 $F_c = [\lambda_0, \lambda_1, \dots, \lambda_{N-1}]$ 来表示样本的结构特征。

1.3 混合结构与时频特征的图卷积神经网络模型

构造 GCNs 有 2 个主要步骤: 输入拓扑图构造和图卷积。将动作数据的结构特征和频谱特征作为图的节点, 并通过确定节点之间的相关性构造图结构。

1.3.1 输入拓扑图构造

将原始动作信号表示为属性图 $G = (X, E, A)$ 。其中, X 为图的输入特征, 其数目为样本总数 M ; E 为图的边集; A 为任意 2 个节点之间连接的

相邻矩阵。

(1)输入特征 \mathbf{X} 。将样本的时频特征 \mathbf{F}_d 和样本的结构特征 \mathbf{F}_c 进行拼接,得到 GCNs 最终的输入特征 \mathbf{X} :

$$\mathbf{X} = \text{hstack}(\mathbf{F}_d, \mathbf{F}_c)。$$
 (8)

式中: $\text{hstack}(\cdot)$ 表示将括号中的元素按水平方向叠加。

(2)边集 \mathbf{A} 。边集最重要的需求就是在相同类别节点之间创建更多的边缘连接的同时减少不同类型之间边缘连接,从而加强同类节点之间的内聚度,降低异类节点之间的耦合度。由于 HVG 构建的样本结构差异性较大,通过样本的结构特征 \mathbf{F}_c 来构建边集难以找到合适的边连接。相同动作在同一个运动周期中的时间序列信号的时频特征具有相似性,本文根据时频特征 \mathbf{F}_d 构建边集:

$$\begin{cases} \mathbf{p}_i = \text{sort}(\mathbf{Dis}_i); \\ \mathbf{S}_i = \{\mathbf{p}_i^m \mid m = 1, 2, \dots, k\}。 \end{cases}$$
 (9)

式中: \mathbf{Dis}_i 为节点 i 特征与其他节点特征之间的欧氏距离; $\text{sort}(\cdot)$ 为升序算子; \mathbf{p}_i 为排序后的 \mathbf{Dis}_i ; \mathbf{S}_i 为 \mathbf{p}_i 中前 k 个数的节点序号集合。最后将节点 i 与 \mathbf{S}_i 中存在的节点之间创建边连接。

(3)相邻矩阵。相邻矩阵 $\mathbf{A} \in \mathbf{R}^{M \times M}$ 的元素为

$$a_{i,j} = \begin{cases} 1, \text{节点 } x_i \text{ 与 } x_j \text{ 存在边连接;} \\ 0, \text{节点 } x_i \text{ 与 } x_j \text{ 不存在边连接。} \end{cases}$$
 (10)

1.3.2 图卷积神经网络

将 $\mathbf{G} = (\mathbf{X}, \mathbf{E}, \mathbf{A})$ 作为 GCN 的输入图形,第 l 层的输出为^[8]

$$\mathbf{H}^{(l)} = \text{ReLU}\left(\tilde{\mathbf{D}} - \frac{1}{2}\tilde{\mathbf{A}}\tilde{\mathbf{D}} - \frac{1}{2}\mathbf{H}^{(l-1)}\mathbf{W}^{(l)}\right)。$$
 (11)

式中: $\text{ReLU}(\cdot)$ 为激活函数; $\mathbf{W}^{(l)}$ 为 GCNs 第 l 层的权重矩阵; $\mathbf{H}^{(l-1)}$ 为第 $l-1$ 层的输出特征且 $\mathbf{H}^{(0)} = \mathbf{X}$; $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$ 意味着在进行图卷积时节点能从自身学习到更多特征; $\tilde{\mathbf{D}}$ 为 $\tilde{\mathbf{A}}$ 的度矩阵,其对角元素 $\tilde{d}_{i,j} = \sum_{i=1}^M \sum_{j=1}^M \tilde{a}_{i,j}$ 。最后将输出的特征输入到 SoftMax 层完成动作分类。

2 实验与结果分析

2.1 实验数据

本文提出的模型在 2 个真实的动作数据集上进行评估。WHARF^[15]数据集在 32 Hz 频率下使用加速度计收集 12 类动作的时序信号。动作包含刷牙、梳头和吃饭等复杂活动。使用长度为 256、重叠率为 50%的固定宽度滑动窗口对数据进行样本划分。DataEgo^[16]数据集在 15 Hz 频率下使用加速度计和

陀螺仪收集 20 类动作的时序信号。动作包含刷牙、梳头和吃饭等复杂活动。使用长度为 256、重叠率为 50%的固定宽度滑动窗口对数据进行样本划分。

2.2 实验环境设置与评测指标

为了评估本文提出的模型,随机选取数据集中 70%的样本作为训练集,30%的样本为测试集。通过准确率 (*Accuracy*),召回率 (*Recall*) 以及 *F1* 分数与文献[4, 17–19]的方法进行比较。原始数据已经转换为输入矩阵 $\mathbf{X}^{M \times 160}$,表示每个样本有 160 个特征。设置隐藏层 $n_{hid1} = 128$ 、二层 GCNs 的输出层 $n_{hid2} = 64$ 。使用学习率为 0.001 的 Adam 优化器, *weight decay* 设置为 10^{-4} ,最大迭代次数为 2 000。所有方法均使用相同的参数运行 10 次,并报告平均结果。所有算法都是由 Python 3.7 用 Pytorch 编写的,并由具有 Intel Xeon E5-2665 CPU 和 32 GB RAM 的服务器处理。

2.3 实验结果

2.3.1 在 2 个数据集上的分类结果对比

表 1 给出了本文提出的 HTC-GCN 模型与其他方法的分类性能对比结果。由表 1 可以看出,本文方法在 DataEgo 数据集上的 3 个常用评价指标都具有优势。同时在 2 个数据集中所提模型在 *F1* 分数上有更好的表现。其中在 WHARF 数据集上,本文模型较对比模型的 *F1* 分数最大提高了 19.58 百分点;在 DataEgo 数据集上,本文模型较对比模型的 *F1* 分数最高提升了 26.44 百分点。这充分说明 GCNs 通过在拓扑结构上聚合相邻节点之间的特征,能实现节点之间信息的有效交流。

表 1 不同方法在 2 个数据集上的分类结果

Table 1 Classification results of different methods on two datasets				%
数据集	方法	Accuracy	Recall	F1
WHARF	文献[4]	67.05	42.88	42.60
	文献[17]	68.73	44.01	43.78
	文献[18]	75.27	58.17	60.49
	本文方法	75.89	60.95	62.18
DataEgo	文献[4]	55.19	43.83	43.00
	文献[17]	58.17	45.71	44.68
	文献[19]	76.14	65.10	64.92
	本文方法	76.97	67.22	69.44

为了更直观地表示 HTC-GCN 网络模型在 WHARF 数据集上的表现,图 3 给出了其识别结果的混淆矩阵图。可以看出,数据集中走路和上下楼、起床和躺、坐下和站起这些高度相似的行为误差率较高,但对于其他大部分行为都能够准确识别其行为类别,证明了本文方法在复杂行为识别上的有效性。

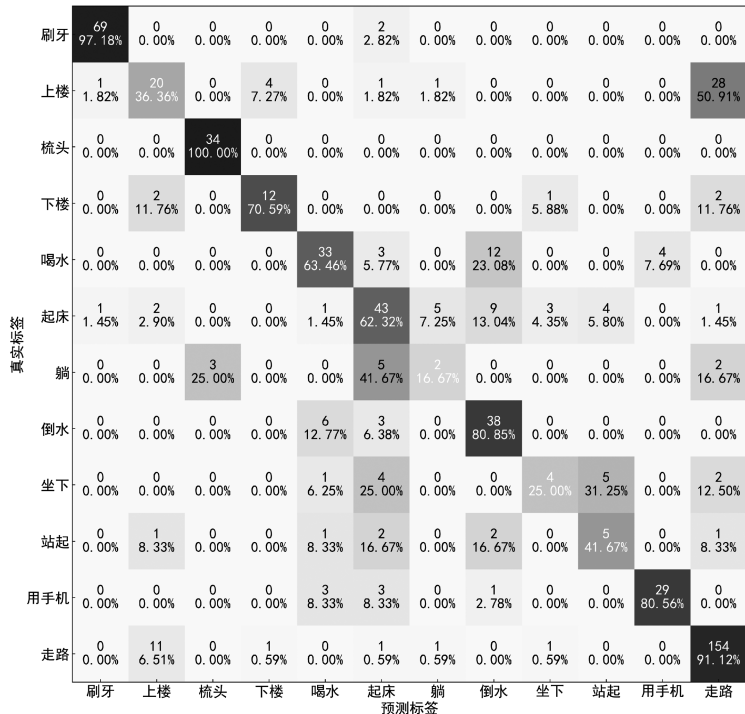


图 3 HTC-GCN 在 WHARF 数据集上准确率混淆矩阵

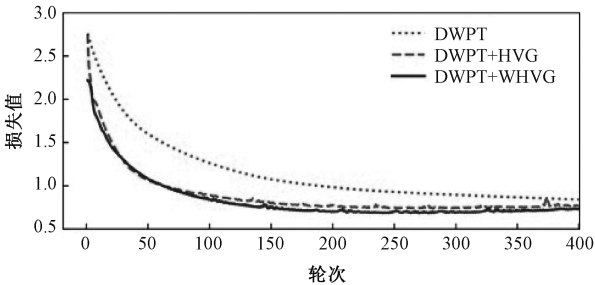
Figure 3 Accuracy confusion matrix of HTC-GCN on WHARF dataset

2.3.2 结构特征对所提模型的性能影响

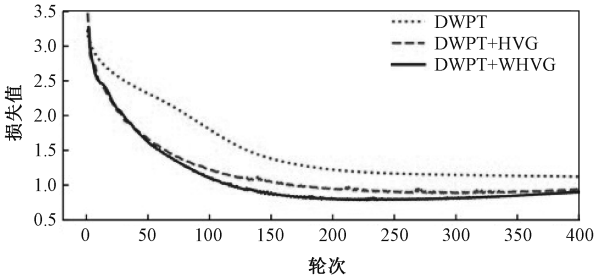
结构特征对本文方法分类结果的影响如表 2 所示,测试过程的损失下降曲线如图 4 所示。其中 DWPT 表示仅使用时频特征对动作进行分类;DWPT+HVG 表示使用不添加约束的结构特征与时频特征融合对动作进行分类;DWPT+WHVG 表示使用添加约束的结构特征与时频特征融合对动作进行分类。显然,在 2 个数据集中,时频特征与结构特征的融合取得了比单一使用时频特征更好的效果。添加上权重约束之后的 DWPT+WHVG 拥有与 DWPT+HVG 相近的准确率和 $F1$ 分数,而损失值的下降速率却快于未加约束的 DWPT+HVG。这充分说明本文通过 HVG 构建时空图,并利用具有额外距离约束的拉普拉斯矩阵表示时空图提取出动作信号结构特征的有效性。

表 2 结构特征对所提方法分类结果的影响

Table 2 Influence of structural characteristics on classification results of the proposed method %			
数据集	方法	Accuracy	$F1$
WHARF	DWPT	68.16	57.79
	DWPT+HVG	75.59	63.09
	DWPT+WHVG	75.89	62.18
DataEgo	DWPT	70.22	63.20
	DWPT+HVG	76.97	68.55
	DWPT+WHVG	76.97	69.44



(a) 在 WHARF 上的损失下降曲线



(b) 在 DataEgo 上的损失下降曲线

图 4 HTC-GCN 模型在 2 个数据集上的损失下降曲线

Figure 4 Loss curves of HTC-GCN model on two datasets

2.3.3 构图方式对所提模型的性能影响

构图方式对 HTC-GCN 分类结果的影响如表 3 所示。其中 TF 表示按时频特征构造 GCNs 输入拓扑图;HVG 表示按加权结构特征构造输入拓扑图;TF+HVG 表示按输入特征构造输入拓扑图。实验表明,单一地通过时频特征构造输入拓扑取得了很好的效果。这不仅说明时频特征构造的输入拓扑图克

服了重采样导致的样本结构差异,而且合适的拓扑构造能帮助 GCNs 更好地完成分类任务。

表 3 构图方式对所提方法分类结果的影响

Table 3 Influence of graph construction on classification results of the proposed method %

数据集	方法	Accuracy	F1
WHARF	TF	75.89	62.18
	HVG	59.64	39.90
	TF+HVG	59.91	39.97
DataEgo	TF	76.97	69.44
	HVG	50.27	37.19
	TF+HVG	50.72	38.12

2.3.4 滑动窗口宽度对所提模型的性能影响

滑动窗口宽度对 HTC-GCN 分类结果的影响如表 4 所示。数据集 WHARF 和 DataEgo 在窗口宽度为 256 时取得最好的 F1 结果。实验表明,窗口宽度并非越大识别效果越好:窗口越大,样本中所含动作周期量变多的同时样本数目减少,削弱了深度模型的学习能力;窗口越小,样本数目增多,样本中所含动作周期量减少,样本特征不足,同样削弱了模型的学习能力。

表 4 滑动窗口宽度对所提方法分类结果的影响

Table 4 Influence of sliding window width on classification results of the proposed method %

数据集	窗口大小	Accuracy	F1
WHARF	24	57.44	49.72
	36	58.34	50.94
	64	60.92	53.68
	120	66.76	58.34
	128	65.86	56.86
	256	75.89	62.18
	512	90.49	57.44
DataEgo	24	69.81	61.49
	36	70.79	63.07
	64	73.24	65.66
	120	76.01	68.22
	128	75.35	67.88
	256	76.97	69.44
	512	67.12	53.14

3 结论

本文在关注可穿戴传感器采样点之间结构信息的同时捕获动作样本之间的潜在关系,探索 GCNs 在人体动作时间序列建模上的可能。在 WHARF 和 DataEgo 数据集上取得的显著改进表明,本文提出的 HTC-GCN 克服了传统深度学习方法的局限性,

初步证明挖掘非结构化时间序列的结构信息是提升可穿戴 HAR 性能的有效途径。基于 GCNs 的 HAR 模型还有很多需要探索的问题,包括节点内部特征和节点关系对分类模型的贡献度等。

参考文献:

[1] HUAN Z, LV S Y, HOU Z J, et al. An evaluation strategy for the symmetry and consistency of lower limb segments during upper limb loading[J]. IEEE Sensors Journal, 2021, 21(5): 6440–6449.

[2] PHAM C, NGUYEN-THAI S, TRAN-QUANG H, et al. SensCapsNet: deep neural network for non-obtrusive sensing based human activity recognition[J]. IEEE Access, 2020, 8: 86934–86946.

[3] ZHU R, XIAO Z L, LI Y, et al. Efficient human activity recognition solving the confusing activities via deep ensemble learning[J]. IEEE Access, 2019, 7: 75490–75499.

[4] CHEN Y Q, XUE Y. A deep learning approach to human activity recognition based on single accelerometer[C]//2015 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE, 2015: 1488–1492.

[5] IGNATOV A. Real-time human activity recognition from accelerometer data using convolutional neural networks[J]. Applied Soft Computing, 2018, 62: 915–922.

[6] WANG L K, LIU R Y. Human activity recognition based on wearable sensor using hierarchical deep LSTM networks[J]. Circuits, Systems, and Signal Processing, 2020, 39(2): 837–856.

[7] SU T T, SUN H Z, MA C M, et al. HDL: hierarchical deep learning model based human activity recognition using smartphone sensors[C]//2019 International Joint Conference on Neural Networks (IJCNN). Piscataway: IEEE, 2019: 1–8.

[8] ZHOU J, CUI G Q, HU S D, et al. Graph neural networks: a review of methods and applications[J]. AI Open, 2020, 1: 57–81.

[9] AHMAD N, CHOW S H C, LEUNG H F. Beyond the gates of Euclidean space: temporal-discrimination-fusions and attention-based graph neural network for human activity recognition[EB/OL]. (2022-06-10) [2023-11-02]. <https://arxiv.org/abs/2206.04855>.

[10] DONNER R V, NORBERT M, YONG Z, et al. Nonlinear time series analysis by means of complex networks[J]. Scientia Sinica Physica, Mechanica & Astronomica, 2020, 50(1): 010509.

[11] ZHU S L, GAN L. Specific emitter identification based on horizontal visibility graph[C]//2017 3rd IEEE International Conference on Computer and Communications

(ICCC). Piscataway: IEEE, 2017: 1328–1332.

[12] LI C Y, MO L F, YAN R Q. Fault diagnosis of rolling bearing based on WHVG and GCN[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 3519811.

[13] 高金峰, 秦瑜瑞, 殷红德. 基于小波包变换和支持向量机的故障选线方法[J]. 郑州大学学报(工学版), 2020, 41(1): 63–69.

GAO J F, QIN Y R, YIN H D. Fault line selection based on wavelet packet transform and support vector machine [J]. Journal of Zhengzhou University (Engineering Science), 2020, 41(1): 63–69.

[14] LOTFOLLAHI-YAGHIN M A, KOOHDARAGH M A. Examining the function of wavelet packet transform (WPT) and continues wavelet transform (CWT) in recognizing the crack specification [J]. KSCE Journal of Civil Engineering, 2011, 15(3): 497–506.

[15] BRUNO B, MASTROGIOVANNI F, SCORBISSA A. Wearable inertial sensors: applications, challenges, and public test benches [J]. IEEE Robotics & Automation Magazine, 2015, 22(3): 116–124.

[16] POSSAS R, CACERES S P, RAMOS F. Egocentric activity recognition on a budget[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 5967–5976.

[17] 郭毅博, 孟文化, 范一鸣, 等. 基于可穿戴传感器数据的人体行为识别数据特征提取方法[J]. 计算机辅助设计与图形学学报, 2021, 33(8): 1246–1253.

GUO Y B, MENG W H, FAN Y M, et al. Wearable sensor data based human behavior recognition: a method of data feature extraction[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(8): 1246–1253.

[18] JORDAO A, TORRES L A B, SCHWARTZ W R. Novel approaches to human activity recognition based on accelerometer data[J]. Signal, Image and Video Processing, 2018, 12(7): 1387–1394.

[19] HUANG Y, YANG X S, GAO J Y, et al. Knowledge-driven egocentric multimodal activity recognition [J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 16(4): 133.

Human Activity Recognition Based on Hybrid Feature Graph Convolutional Neural Network

LI Zhixin¹, SHANG Fanqi¹, HUAN Zhan¹, CHEN Ying¹, LIANG Jiuzhen²

(1. School of Microelectronics and Control Engineering, Changzhou University, Changzhou 213000, China; 2. School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou 213000, China)

Abstract: The existing methods for human activity recognition using wearable sensors could not capture the structural information between the sampling points of time series effectively and might ignore the potential connections between samples. To address this issue, a graph convolutional neural network model with hybrid time-frequency and structural characteristics was proposed for human activity recognition. Firstly, the time-frequency characteristics of the original signal were obtained by wavelet-packet transform, and the spatio-temporal graph was further constructed to extract the structural characteristics to identify the dynamic characteristics between the sampling points. The distance constraint was added to the structural characteristics to weaken the influence of long-distance neighbors on the central node on the spatio-temporal graph. Considering that the extraction of structural characteristics was greatly affected by the topological relationship of the spatio-temporal graph. The time-frequency characteristics of the samples to construct the input topology of the graph convolutional neural network were selected, and the time-frequency and structural characteristics were combined as the input features of the network. Finally, the input features propagated along the input topology to obtain the final classification result. To evaluate the performance of the proposed model, experiments were conducted on the WHARF and DataEgo datasets. Results in terms of *F1* scores indicated that the proposed model outperformed existing convolutional neural network-based methods, achieving a maximum improvement of 19.58 percentage points on the WHARF dataset and 26.44 percentage points on the DataEgo dataset. It demonstrated that the proposed model could effectively enhance the capability of activity recognition by exploiting dynamic characteristics.

Keywords: graph convolutional neural networks; wearable device; human activity recognition; spatio-temporal graph; feature extraction