

马鹏飞, 卫芬, 沈意平, 等. 人体行为多尺度卷积双向门控循环识别模型[J]. 湖南科技大学学报(自然科学版), 2024, 39(4): 98-108. doi:10.13582/j.cnki.1672-9102.2024.04.012

MA P F, WEI F, SHEN Y P, et al. Human Activity Recognition Model Based on Multi-Scale Convolutional-Bidirectional Gated Recurrent Unit [J]. Journal of Hunan University of Science and Technology (Natural Science Edition), 2024, 39(4): 98-108. doi:10.13582/j.cnki.1672-9102.2024.04.012

人体行为多尺度卷积双向门控循环识别模型

马鹏飞¹, 卫芬^{2,3}, 沈意平^{1*}, 吕泽强³

(1. 湖南科技大学 机械设备健康维护湖南省重点实验室, 湖南 湘潭 411201;

2. 国防科技大学 装备综合保障技术重点实验室, 湖南 长沙 410073;

3. 湖南科技大学 机电工程学院, 湖南 湘潭 411201)

摘要:为实现深度挖掘人体行为传感数据中的前后关联信息,进而达到精确识别人体行为模式的目的,针对传统卷积神经网络无法充分利用行为特征之间蕴藏的时序信息,提出一种基于注意力机制的多尺度卷积神经网络(Multi-scale Attention Convolutional Neural Network, MACNN)与双向门控循环单元(Bidirectional Gated Recurrent Unit, BiGRU)的人体行为识别模型.提出一种多尺度卷积结构,拓宽网络的宽度并实现不同维度特征的提取和降维;引入双向门控循环单元学习信号在时间维度上的关联特征;通过在多尺度 CNN 中加入卷积注意力模块,深度挖掘人体行为信号中的关键信息特征,实现人体行为的高准确度识别.采用 UCI 人体行为识别公共数据集对所提方法进行试验验证,结果表明:所提方法能够准确地实现对 6 种典型人体日常行为的分类识别,且相比于传统的单尺度 CNN 识别模型,所提出的 MACNN-BiGRU 模型的准确率提高 5% 以上,达到 98.40%.

关键词:人体行为识别;多尺度卷积神经网络;注意力机制;双向门控循环单元

中图分类号:TP391

文献标志码:A

文章编号:1672-9102(2024)04-0098-11

Human Activity Recognition Model Based on Multi-Scale Convolutional-Bidirectional Gated Recurrent Unit

MA Pengfei¹, WEI Fen^{2,3}, SHEN Yiping¹, LYU Zeqiang³

(1. Hunan Provincial Key Laboratory of Health Maintenance for Mechanical Equipment, Hunan University of Science and Technology, Xiangtan 411201, China;

2. Laboratory of Science and Technology on Integrated Logistics Support, National University of Defense Technology, Changsha 410073, China;

3. School of Mechanical and Electrical Engineering, Hunan University of Science and Technology, Xiangtan 411201, China)

Abstract: In order to realize the deep mining of the backward and forward relevant information in human activity sensing data, and then achieve the purpose of accurately identifying human behavior patterns, meanwhile aiming at the issue that the traditional convolutional neural network cannot make full use of the time-series information contained in the behavioral features, a human activity recognition model based on Multi-scale Attention Convolutional Neural Network with attention mechanism (MACNN) and Bidirectional Gated Recurrent Unit (BiGRU) is proposed. Firstly, a multi-scale convolution structure is proposed to broaden the width of the

收稿日期:2022-06-27

修改日期:2024-09-25

基金项目:湖南创新型省份建设专项经费资助(2020RC3049);湖南省教育厅科学研究项目资助(20C0775;21A0310);湖南省自然科学基金青年项目资助(2023JJ40289)

* 通信作者, E-mail: yiping1011@163.com

network and realize the extraction and dimensionality reduction of different dimensional features. Secondly, the bi-directional gated recurrent unit is introduced to learn the correlation characteristics in the time dimension of the signal, and thirdly, by adding convolution attention module into the aforementioned MCNN, the key information features in human activity signals are deeply mined, and ultimately high-precision human activity recognition is realized. The public data set of UCI-HAR is utilized to verify the validity of the proposed recognition model. The results indicate that the proposed model can accurately classify and recognize six kinds of typical human activity. Compared with the traditional single-scale CNN, the accuracy of the proposed model is improved by more than 5%, and the recognition accuracy reaches 98.4%.

Keywords: human activity recognition; multi-scale convolutional neural networks; attention mechanism; bidirectional gated recurrent unit

人体行为识别技术(Human Activity Recognition, HAR)通常借助可穿戴设备中的传感器实现对人体行为数据的采集,结合智能学习方法进行人体行为的分类识别^[1-3].传统的人体行为识别方法主要有支持向量机^[4]、K-近邻^[5]、随机森林^[6]和隐马尔科夫模型^[7]等,上述方法均建立在人工特征提取的基础上,且强依赖于丰富的先验知识及相关专家工程经验.

深度学习算法具有无需人工构造特征即可实现深度挖掘数据本身隐含的内在特征的特点,目前已被广泛应用于各类模式识别领域.作为深度学习框架下的一种典型模型,卷积神经网络(Convolutional Neural Network, CNN)通过自动特征提取以及权值共享的方式快速训练模型.由于人体行为数据通常表现为具有前后关联特点的时序信号,在时间维度上蕴含丰富的价值信息,单一的CNN模型无法深度挖掘特征之间蕴藏的时序信息,而CNN与循环神经网络(Recurrent Neural Network, RNN)的结合,能够充分发挥CNN的特征提取能力和RNN对时间序列分析的特点,研究者们通常采用两者的组合模型完成人体行为的准确识别^[8-9].HOCHREITER等^[10]提出的长短期记忆网络(Long Short-Term Memory, LSTM)是一种改进的RNN,解决了传统RNN中梯度消失难以训练网络的问题,具备更强的时间序列信号处理能力;杨万鹏等^[11-12]提出CNN与LSTM相结合的方法,该方法在人体行为数据集上可以达到96.11%的识别准确率.考虑到LSTM的结构复杂、收敛速度较慢等问题,CHUNG等^[13]于2014年提出门控循环单元(Gated Recurrent Unit, GRU),旨在提高LSTM准确率的同时优化网络结构,提高收敛速度;王震宇等^[14]将CNN和GRU融合作为学习模型,执行自动特征提取和人体行为分类.然而,GRU只考虑先前的人体行为信息,忽略了未来状态对人体行为识别的影响,无法充分提取人体行为数据之间潜在的前后关联特征^[15],因此,引入双向门控循环单元(Bidirectional Gated Recurrent Unit, BiGRU)能够更加深入地挖掘信号中的时间序列信息,有效增强前后信息之间的关联性.

目前,CNN在人体行为识别方面取得了明显成效,但存在以下问题:(1)由于人体行为数据具有复杂的多尺度特点,随着网络层数的增加,网络参数个数呈指数级增加,极易发生过拟合现象,无法充分利用行为特征之间蕴藏的时序信息;(2)现有的深度学习方法存在梯度消失难以训练网络的问题;(3)训练迭代过程中包含大量参数,导致网络收敛速度缓慢.因此,本文提出一种基于注意力机制的多尺度卷积神经网络与双向门控循环单元的人体行为识别模型,简称为MACNN-BiGRU,其主要贡献如下:

1)通过多尺度的卷积核挖掘与人体行为变化密切相关的局部空间特征信息,旨在有效减少网络参数和计算量,提高网络效率的同时有效防止过拟合现象.

2)采用BiGRU充分利用信号间的序列关系学习时间维度上的信息特征,在不损失模型精度的前提下提高模型的速度.

3)在多尺度卷积层中加入卷积注意力模块(Convolutional Block Attention Module, CBAM),使其在空间维度和通道维度建立联系,加强模型对于关键信息的提取能力,获得更加准确的结果.

1 人体行为识别模型的构建

1.1 双向门控循环单元

门控循环单元(GRU)作为LSTM衍生出的一种变体,能够更好地捕捉序列中的时间关联特性^[16-17].

GRU 将 LSTM 中的输入门和遗忘门合并成一个更新门,能够保留 LSTM 长期记忆的优点,减少网络的参数,提升模型的训练速度,其结构如图 1 所示。

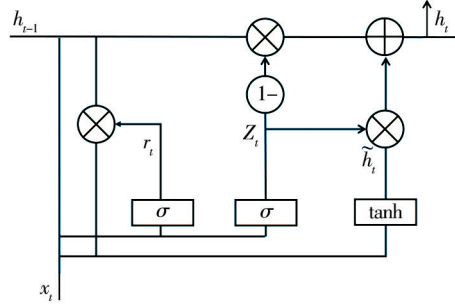


图 1 GRU 结构

GRU 主要包括重置门 r_t 和更新门 z_t (t 表示当前时刻)。通过更新门 z_t 将指定相关信息保留至下一个状态,重置门 r_t 将先前信息状态与新的输入信息相结合。在处理人体行为时间序列数据时,GRU 能有效学习其中的长期关联特性,其具体计算过程为

$$z_t = \sigma(\mathbf{w}_z x_t + \mathbf{w}_z h_{t-1} + b_z); \quad (1)$$

$$r_t = \sigma(\mathbf{w}_r x_t + \mathbf{w}_r h_{t-1} + b_r); \quad (2)$$

$$\tilde{h}_t = \tanh(\mathbf{w}_h x_t + r_t \otimes \mathbf{w}_h h_{t-1}); \quad (3)$$

$$h_t = z_t \otimes h_{t-1} + (1 - z_t) \otimes \tilde{h}_t. \quad (4)$$

式中: $\sigma(\cdot)$ 与 $\tanh(\cdot)$ 分别为 Sigmoid 与双曲正切激活函数; x_t 为输入数据; h_{t-1} 为前一状态的输出数据; $\mathbf{w}_z, \mathbf{w}_r$ 和 $\mathbf{w}_h$ 分别为更新门、重置门以及输出待定值的权重矩阵; b_z, b_r 为更新门、重置门输出的偏置,这些偏置项在计算过程中起到调整门控单元输出值的作用,确保输出值在合适的范围内,从而更好地控制信息的流动。

双向门控循环单元 (BiGRU) 由 2 个单向 GRU 构成^[18], 每个 GRU 分别沿着时间正序和逆序的方向对输入序列进行自动特征提取,然后通过融合层对特征进行加权处理,最后合并为输出序列。BiGRU 的结构如图 2 所示。

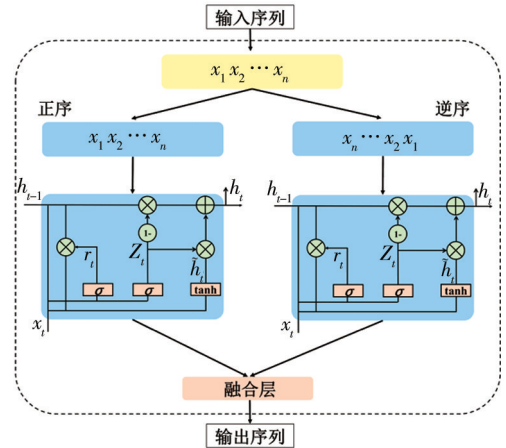


图 2 BiGRU 结构

1.2 注意力机制

卷积注意力模块是以人类视觉系统的处理机制为原理,对数据特征分配不同的权重,使模型更关注于重要信息的提取,从而提高模型的特征学习能力^[19]。CBAM 包含通道注意力模块和空间注意力模块,结构如图 3 所示,整体计算过程为

$$F' = M_c(F) \otimes F; \quad (5)$$

$$F'' = M_s(F') \otimes F'. \quad (6)$$

式中: $M_c(\cdot)$ 为通道注意力模块; $M_s(\cdot)$ 为空间注意力模块; F' 为通道注意力模块输出结果; F'' 为 CBAM 输出结果。

卷积层提取的特征图 $F (F \in \mathbf{R}^{H \times W \times C})$ 为输入特征,其中 H, W, C 分别为特征图的高、宽和通道数。首先通过通道注意力模块,计算出输入特征各个通道的权重。特征图 F 分别采用全局最大池化 (Max Pooling) 和平均池化 (Average Pooling) 得到 2 个 $1 \times 1 \times C$ 的全局最大池化特征 $F_{\max}^C$ 和全局平均池化特征 F_{avg}^C , 再通过多层感知机 (Multilayer Perception, MLP) 得到 2 个特征向量,再将这 2 个特征向量的元素逐个相加累积,经 Sigmoid 函数激活,产生通道注意力特征图 $M_c(F)$, 其中 W_0 和 W_1 为 MLP 的权重, σ 为激活函数。

$$M_c(F) = \sigma \{ \text{MLP} [\text{MaxPool}(F)] + \text{MLP} [\text{AvgPool}(F)] \} = \sigma \{ W_1 [W_0 (F_{\max}^C)] + W_1 [W_0 (F_{\text{avg}}^C)] \}. \quad (7)$$

$M_c(F)$ 与输入特征 F 的元素相乘,得到通道注意力模块的输出特征图,记为 F' ,同时作为空间注意力模块的输入.空间注意力机制提取特征图上各空间区域之间的联系, F' 通过最大池化和平均池化层分别得到通道维度上的全局最大池化特征 $F_{\max}^S$ 和全局平均池化特征 F_{avg}^S ,经过卷积和激活函数得到空间注意力特征图 $M_s(F')$,过程如式(8)所示.

$$M_s(F') = \sigma[\text{Conv}(|\text{MaxPool}(F')|; |\text{AvgPool}(F')|)] = \sigma[\text{Conv}(F_{\max}^S; F_{\text{avg}}^S)]. \quad (8)$$

式中:Conv 为核尺寸为 7×1 的卷积操作.

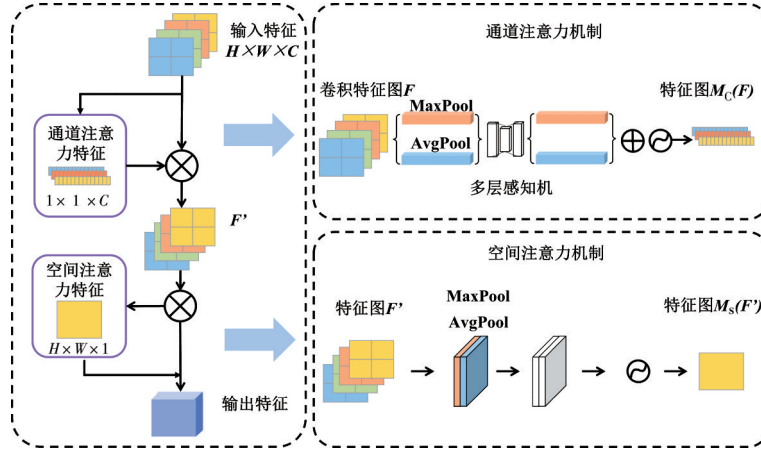


图3 CBAM 结构

通过计算不同尺度特征的通道注意力特征图 F' 和空间注意力特征图 F'' , 提高人体行为特征的表征能力,捕捉特征图上各空间区域之间的联系并抑制复杂背景噪声的影响.

1.3 基于 MACNN-BiGRU 的人体行为识别模型

本文针对传统卷积神经网络进行人体行为识别时会产生部分特征丢失且无法充分利用特征之间时序信息的不足,提出一种基于 MACNN-BiGRU 的人体行为识别模型,模型结构如图4所示,包括多尺度 CNN,双向门控循环单元和卷积注意力模块.

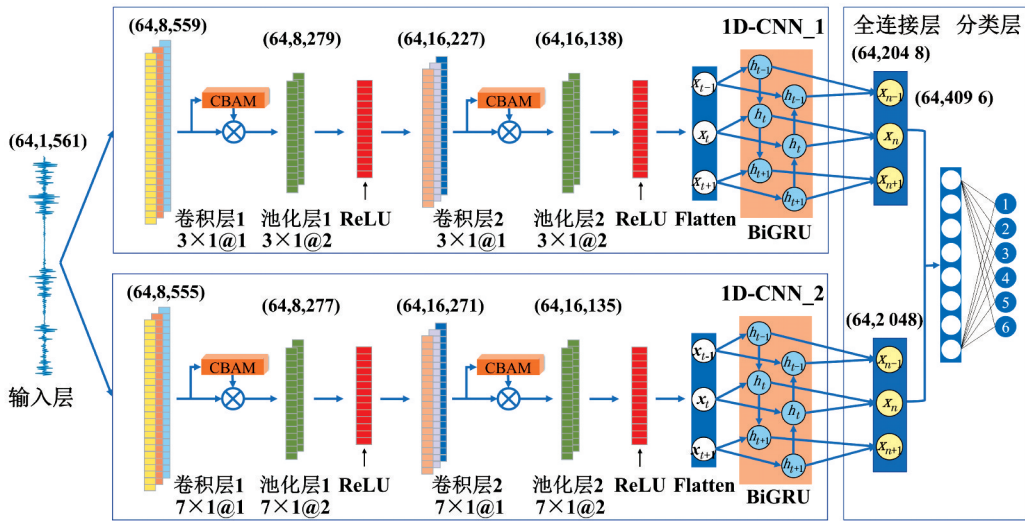


图4 MACNN-BiGRU 结构

基于人体行为数据,搭建多尺度特征提取层,每一个分支均采用不同尺度的卷积核构成,分别记为 CNN_1 和 CNN_2,自上而下,卷积核大小分别为 3×1 , 7×1 .第一阶段为初步特征提取,将一维数据输入第一层卷积,采用 64 个卷积核为 8,尺度为 3×1 ,步长为 1 的卷积核进行特征提取;然后在多尺度卷积层引入注意力模块,对通道与空间的重要特征赋予较大的权重,提高模型提取关键行为特征的能力;通过池化层对提取的特征进行降采样.第二阶段将得到的特征图输入多尺度卷积模块,拓宽网络的宽度并提取到了多

尺度和更加丰富的特征.第三阶段引入双向门控循环单元进行时间序列分析,可以获取更多结构内在的空间、时间相关的特征信息,最后通过全连接层对不同分支的特征矩阵进行融合,再传入分类层进行人体日常行为识别.

1.4 基于 MACNN-BiGRU 模型的人体行为识别实施流程

针对人体行为数据的识别,本文提出一种基于 MACNN-BiGRU 的人体行为识别模型,具体实施步骤如下:

- 1) 获取人体行为数据集,通过滤波去噪、滑动窗口分割等方法进行数据预处理.
- 2) 对预处理的数据集进行随机划分,得到训练集和测试集;构建 MACNN-BiGRU 模型,并对网络模型参数初始化,通过反向传播参数更新权重,完成模型训练.
- 3) 根据测试集样本的预测标签与真实标签进行对比,获取模型故障识别的准确率,同时采用混淆矩阵(Confusion Matrix)^[20]、t-分布邻域嵌入算法(t-SNE)^[21]对测试集分类结果可视化,以此衡量诊断模型的准确性和鲁棒性,具体实施流程如图5所示.

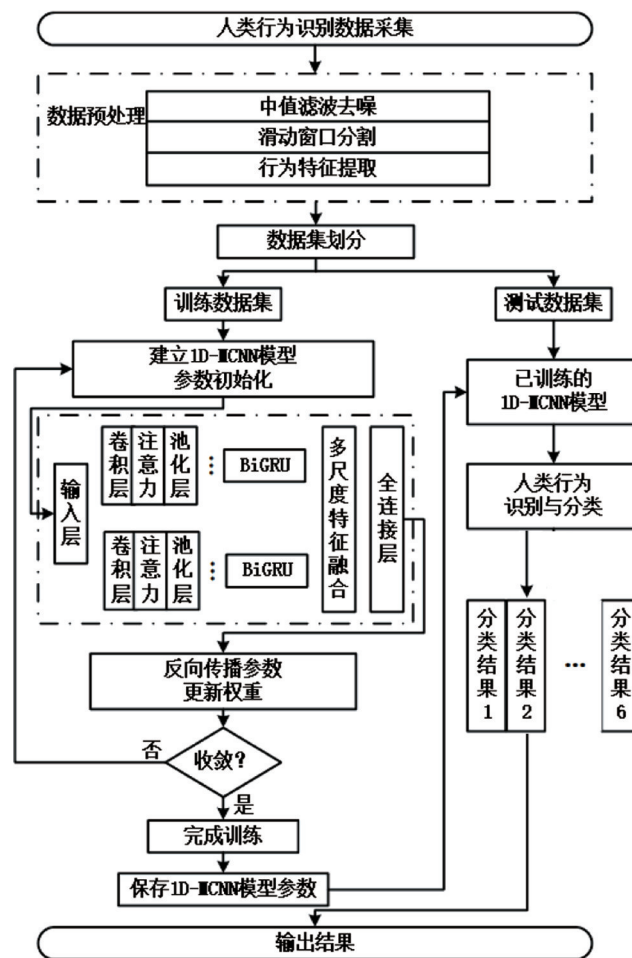


图5 MACNN-BiGRU 识别流程

2 基于 UCI-HAR 数据集的试验验证

试验使用的人体行为识别数据来自机器学习网站 UCI 中的开源 HAR 数据集^[22],该数据集是采用 Samsung Galaxy SII 手机内置的三轴加速度计和三轴陀螺仪 2 种惯性传感器进行 X 轴、Y 轴、Z 轴的数据采集,包括 30 个年龄在 19~48 岁的志愿者完成 3 种静态行为(坐、站立、躺下)和 3 种动态行为(步行、上楼、下楼)数据.传感器的采样频率为 50 Hz.将不同测试者随机分成 2 组,随机挑选 70% 的测试者作为训练集数据,剩余 30% 为测试集数据,表 1 为各行为样本所占比例.

表 1 各行为样本分布

行为模式	总体样本数	分布比例/%	标签编号
步行	1 722	16.72	0
下楼	1 544	14.99	1
上楼	1 406	13.66	2
坐	1 777	17.24	3
站立	1 906	18.51	4
躺下	1 944	18.88	5

2.1 模型训练过程

本文对采集信号进行预处理,通过截止频率为 20 Hz 的三阶低通 Butterworth 滤波器降低噪声.采用宽度为 128,重叠率为 50%的滑动窗口对信号进行分割,每次采样可获得三轴加速度信号以及三轴角速度信号.将采样数据进行拼接融合,在时域信号中提取 561 个包含人体行为活动的时间频率特征向量,共计 10 299 个样本进行训练与验证.提取其中一位测试者的三轴加速度和角速度信号如图 6 所示,人体的倾斜角不同,对应重力加速度各分量的值也不同,如图 6a 所示.角速度数据则呈现明显的周期性变化,数据在正负范围内均有出现.由于人体行为数据之间蕴藏的时序信息,采用深度学习方法可以直接将原始人体行为数据作为网络模型的输入,避免丢失重要的时域信息,同时深度挖掘人体行为信号中的关键信息特征,实现人体行为模式的高精度识别.

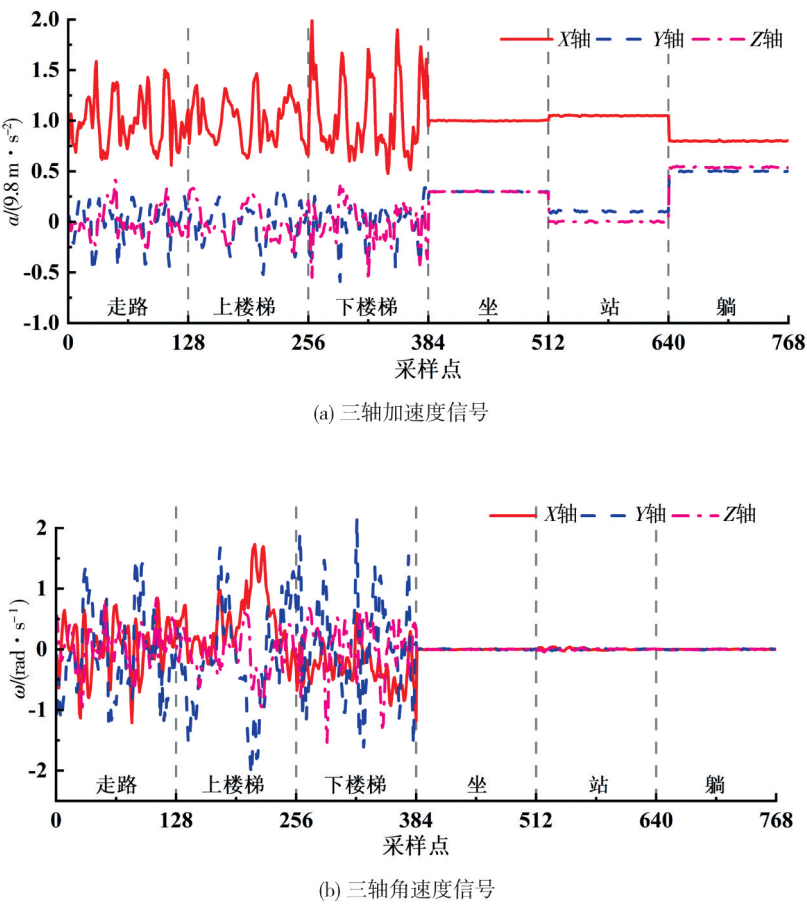


图 6 人体运动数据分析

利用搭建好的 MACNN-BiGRU 模型对上述 HAR 数据集进行训练,使用 Dropout 层防止模型过拟合,Dropout 比率为 0.2,初始学习率为 0.01,hidden_size 为 64,BiGRU 网络层数为 2,采样用交叉熵作为损失函

数^[23].本文模型是在基于 Python 的科学计算包 PyTorch 中进行搭建,训练和测试的硬件环境为 Intel i5-7300H,借助 GPU 显卡硬件提升计算性能,加速数据处理与模型训练.MACNN-BiGRU 模型的参数如表 2 所示.

表 2 MACNN-BiGRU 模型参数

网络层	名称	卷积核大小@步长	卷积核数	输出尺寸
CNN_1	Input			(64,1,561)
	Conv1	3×1@1	8	(64,8,559)
	CBAM	3×1@1	8	(64,8,559)
	Max_Pool1	3×1@2	8	(64,8,279)
	Conv2	3×1@1	16	(64,16,277)
	CBAM	3×1@1	16	(64,16,277)
	Max_Pool2	3×1@2	16	(64,16,138)
	BiGRU			(64,16,138)
	FC1			(64,2 048)
CNN_2	Conv1	7×1@1	8	(64,8,555)
	CBAM	7×1@1	8	(64,8,555)
	Max_Pool1	7×1@2	8	(64,8,277)
	Conv2	7×1@1	16	(64,16,271)
	CBAM	7×1@1	16	(64,16,271)
	Max_Pool2	7×1@2	16	(64,16,135)
	BiGRU	7×1@1	16	(64,16,135)
融合层	FC2			(64,2 048)
	FC			(64,4 096)
	Output			(64,6)

为了更好地描述 MACNN-BiGRU 模型的训练效果,针对模型在训练过程中每次迭代的训练集、测试集的准确率曲线与损失曲线的趋势进行深入分析,结果如图 7 所示.随着迭代次数的增加,所提出的 MACNN-BiGRU 模型的准确率总体呈上升趋势,损失率相应地呈下降趋势,在迭代 20 次后,损失率下降变缓,准确率曲线的斜率趋于 0.当 Epoch 为 64 左右时,其准确率与损失率均轻微浮动,模型趋于稳定,达到了良好的收敛效果.迭代 100 次时其识别准确率达到 98.4%左右,由此可知,所提出的 MACNN-BiGRU 模型具有良好的识别效果.

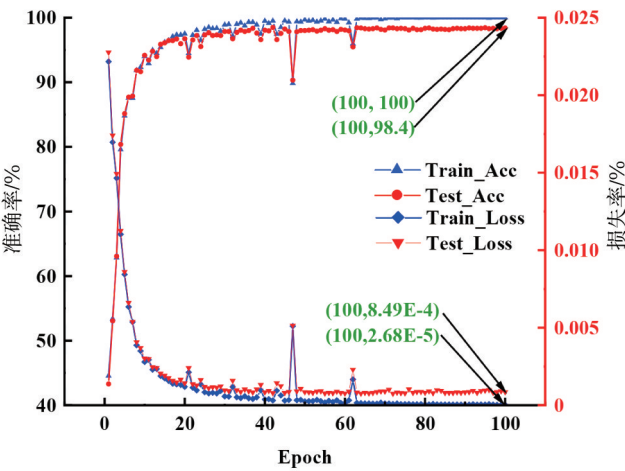


图 7 MACNN-BiGRU 模型准确率曲线及损失曲线

2.2 分类性能影响因素分析

本节讨论不同尺度特征提取与注意力机制的引入对分类性能的影响.首先对 MACNN-BiGRU 与传统

的 CNN-BiGRU 模型在相同的分类器和数据集条件下进行对比分析,采用准确率和 F1 分数指标来评价模型性能,具体公式为

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}; \quad (9)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}; \quad (10)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}; \quad (11)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (12)$$

式中:TP,FP 为真阴性和假阳性的数量;TN,FN 为真阴性和假阴性的数量.

表 3 对比分析了 MACNN-BiGRU 和 CNN-BiGRU 这 2 种模型在 6 种不同行为模式上的独立识别效果.由表 3 可知:单一尺度的 CNN-BiGRU 模型的平均准确率为 96.10%,而所提出的 MACNN-BiGRU 模型的平均准确率达到 98.40%.

表 3 基于不同尺度模型的行为模式识别结果对比

行为模式	CNN-BiGRU		MACNN-BiGRU	
	准确率/%	F1 分数/%	准确率/%	F1 分数/%
步行	98	98	99	98
下楼	97	97	98	99
上楼	95	97	100	99
坐	96	95	96	96
站立	94	91	97	98
躺下	97	99	100	100

2 种不同模型的混淆矩阵能够更加直观地对分类结果的正确性进行量化分析,UCI-HAR 混淆矩阵结果如图 8 所示,其对角元素表示识别准确率.由图 8a 中可以看出:单一尺度的 CNN-BiGRU 模型对上楼、下楼状态以及坐、站立状态的混淆程度较大,有 6%的样本分类错误.而图 8b 中 MACNN-BiGRU 模型区分坐和站立状态时仅有 3%~4%的样本被误分类,其他 4 种活动状态(步行、上楼、下楼和躺下)分类结果的误差较小,相对于 CNN-BiGRU 而言,MACNN-BiGRU 模型在测试过程中不仅整体识别率更高,而且能够更好地区分不同的人体行为状态.

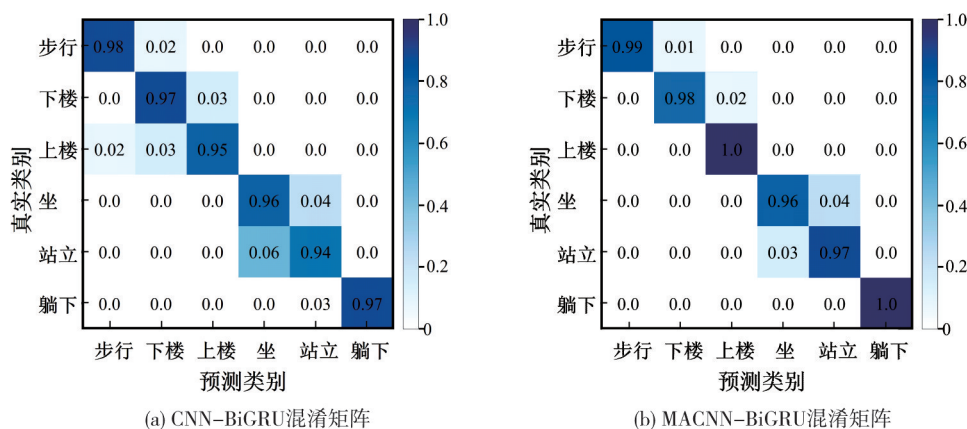


图 8 UCI-HAR 混淆矩阵结果

为了更深入地了解 MACNN-BiGRU 模型的内在学习机理,采用 t-SNE 降维技术^[24]对模型中 CNN_1 的输入层、卷积层、BiGRU 以及最后的输出层进行特征数据的可视化处理,结果如图 9 所示.

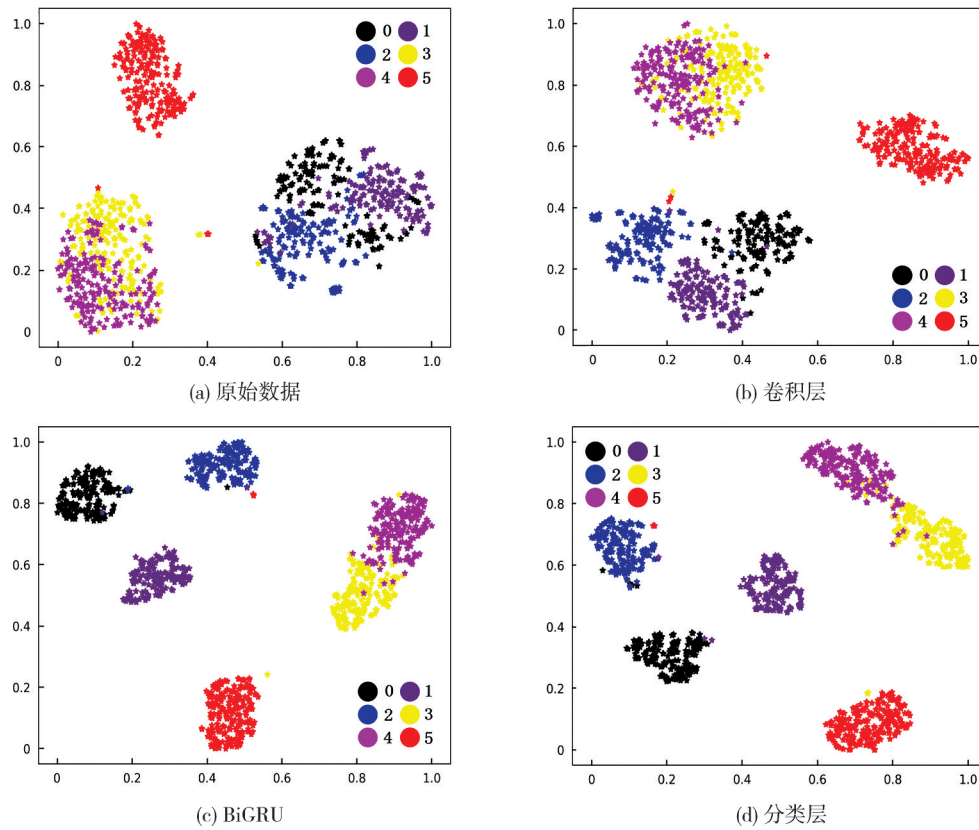


图 9 t-SNE 可视化

由图 9 可知:原始数据中不同行为相对混叠,可分性较差,经过卷积层、BiGRU 的自适应学习后,不同行为模式的区分度逐渐提高.所提出的 MACNN-BiGRU 模型可以对人体行为进行准确地分类识别.为了证明注意力机制引入的优越性,将多尺度卷积神经网络与 BiGRU 组合的 MCNN-BiGRU 模型与在多尺度卷积层中添加 CBAM 的 MACNN-BiGRU 模型进行对比试验,迭代 100 次时各模型测试集的准确率曲线与损失曲线如图 10 所示.

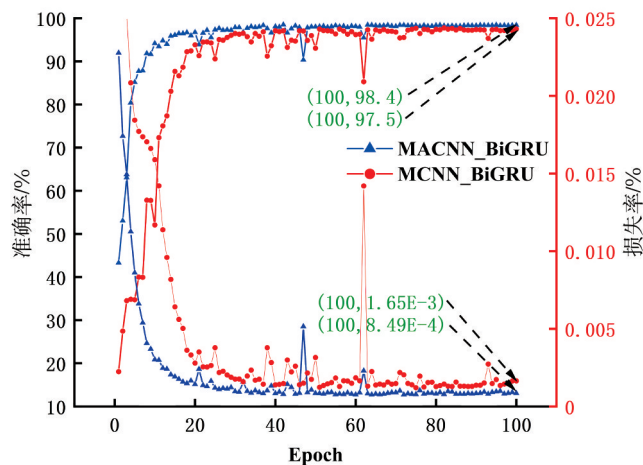


图 10 注意力机制下模型准确率与损失曲线对比

由图 10 可知:引入卷积注意力机制以后,MACNN-BiGRU 模型的准确率为 98.4%,与 MCNN-BiGRU 模型相比,准确率提高 0.9 个百分点,同时曲线振荡显著减少,经过 15 次迭代后准确率基本稳定,说明模型的收敛速度快且稳定.结果表明:在原始时域信号输入以及其他网络结构均相同的前提下,注意力机制的加入使本文提出的 MACNN-BiGRU 模型对时序信号的重要特征赋予较大的权值,有效提高了模型对人

体行为数据的识别精度.

2.3 不同模型识别结果分析

将本文模型与使用 UCI-HAR 公共数据集的相关文献提出的模型进行性能对比,同样采用准确率和 F1 分数指标对不同模型在 6 种不同行为模式上的综合识别效果进行分析评价.对比模型主要包括传统的 CNN 模型、LSTM 网络、DeepConvGRU 模型^[23]、CNN-LSTMs 模型^[12]、CNN-BiGRU 模型和无注意力机制的 MCNN-BiGRU 模型,对比分析结果如表 4 所示.

表 4 基于 UCI-HAR 数据集的不同方法比较结果

方法	准确率/%	F1 分数/%
CNN	92.70	92.90
LSTM	88.99	89.01
DeepConvGRU	93.85	—
CNN-LSTMs	95.93	—
CNN- BiGRU	97.13	97.30
MCNN-BiGRU	97.50	97.62
MACNN-BiGRU	98.40	98.33

由表 4 可知:单一的深度学习方法如 CNN 和 RNN 的准确率明显低于 2 种方法的结合;而在 CNN 和 RNN 结合的方法中,由于 BiGRU 具有沿时间正序和逆序的方向对输入序列进行处理的能力,相比于传统的 LSTM,CNN-BiGRU 提取的特征丰富且收敛速度快;对比有无注意力机制的引入,也可得出本文所提出的 MACNN-BiGRU 模型的性能明显优于无注意力机制的 MCNN-BiGRU 模型.综上可知,相比于现有的识别模型,所提出的 MACNN-BiGRU 模型在人体行为识别中取得了最佳的识别效果.

3 结论

1)通过多尺度卷积神经网络挖掘与人体行为变化密切相关的局部空间特征信息,在 MCNN 的卷积层中加入 CBAM,使其在空间维度和通道维度上建立联系,加强模型对于关键信息的提取能力,大大提高了模型的计算速度与识别性能.

2)引入 BiGRU 充分挖掘信号中的时间序列信息,增强前后信息之间的联系,充分融合多尺度卷积神经网络和 BiGRU,使模型具备空间维度和时间维度的特征信息学习能力,有效地避免了传统 RNN 模型中梯度消失的不足.

3)所提出的模型识别的准确率可达到 98.40%,相比于现有的传统模型,其在人体行为识别中表现出最佳的识别性能.

参考文献:

- [1] LIU Z G, WANG L, YIN Z Y, et al. Task-driven joint dictionary learning model for multi-view human action recognition[J]. Digital Signal Processing, 2022, 126: 103487.
- [2] SLADE S, ZHANG L, YU Y H, et al. An evolving ensemble model of multi-stream convolutional neural networks for human action recognition in still images[J]. Neural Computing and Applications, 2022, 34(11): 9205–9231.
- [3] SINGH P K, KUNDU S, ADHIKARY T, et al. Progress of human action recognition research in the last ten years: a comprehensive survey[J]. Archives of Computational Methods in Engineering, 2022, 29(4): 2309–2349.
- [4] GARCIA-GONZALEZ D, RIVERO D, FERNANDEZ-BLANCO E, et al. A public domain dataset for real-life human activity recognition using smartphone sensors[J]. Sensors, 2020, 20(8): 2200.
- [5] PATRO S G K, MISHRA B K, PANDA S K, et al. A hybrid action-related K-nearest neighbour (HAR-KNN) approach for recommendation systems[J]. IEEE Access, 2020, 8: 90978–90991.

- [6] MAO X J, PENG L H, WANG Z L. Nonparametric feature selection by random forests and deep neural networks[J]. Computational Statistics & Data Analysis, 2022, 170: 107436.
- [7] SIDDIQI M H, ALRUWAILI M, ALI A, et al. Human activity recognition using Gaussian mixture hidden conditional random fields[J]. Computational Intelligence and Neuroscience, 2019, 2019: 8590560.
- [8] MAITRE J, BOUCHARD K, GABOURY S. Alternative deep learning architectures for feature-level fusion in human activity recognition[J]. Mobile Networks and Applications, 2021, 26(5): 2076–2086.
- [9] 匡晓华, 何军, 胡昭华, 等. 面向人体行为识别的深度特征学习方法比较[J]. 计算机应用研究, 2018, 35(9): 2815–2817.
- [10] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735–1780.
- [11] 杨万鹏, 李擎, 雷明. 人体动作识别的特征级融合 LSTM-CNN 方法研究[J]. 电子测量技术, 2021, 44(17): 173–180.
- [12] 陈飞, 程合彬, 王伟光. 基于 CNN-LSTMs 混合模型的人体行为识别方法[J]. 信息技术与信息化, 2019(4): 32–34.
- [13] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. ArXiv e-Prints, 2014: arXiv: 1412.3555.
- [14] 王震宇, 张雷. 基于深度卷积和门控循环神经网络的传感器运动识别[J]. 电子测量与仪器学报, 2020, 34(1): 1–9.
- [15] CHEN Y X, JIANG J, QIN X J, et al. Quantitative monitoring of bolt looseness using multichannel piezoelectric active sensing and CBAM-based convolutional neural network[J]. Frontiers in Materials, 2021, 8: 677642.
- [16] 高越, 傅湘玲, 欧阳天雄, 等. 基于时空自适应图卷积神经网络的脑电信号情绪识别[J]. 计算机科学, 2022, 49(4): 30–36.
- [17] 宫文峰, 陈辉, 张泽辉, 等. 基于改进卷积神经网络的滚动轴承智能故障诊断研究[J]. 振动工程学报, 2020, 33(2): 400–413.
- [18] TAO Q, LIU F, LI Y, et al. Air pollution forecasting using a deep learning model based on 1D convnets and bidirectional GRU[J]. IEEE Access, 2019, 7: 76690–76698.
- [19] FU H X, SONG G Q, WANG Y C. Improved YOLOv4 marine target detection combined with CBAM[J]. Symmetry, 2021, 13(4): 623.
- [20] LIU X F, LIU W, HUANG H S, et al. An improved confusion matrix for fusing multiple K-SVD classifiers[J]. Knowledge and Information Systems, 2022, 64(3): 703–722.
- [21] JIANG M K, JIANG H, ZHAO Y, et al. A novel short-term power load forecasting method based on TSNE-EEMD-LSTM[J]. International Transactions on Electrical Energy Systems, 2022, 2022: 4802633.
- [22] ANGUITA D, GHIO A, ONETO L, et al. A public domain dataset for human activity recognition using smartphones[J]. ESANN 2013 Proceedings, 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2013: 437–442.
- [23] 李君君, 张彬彬, 江朝晖. 基于形变卷积神经网络的行为识别[J]. 智能计算机与应用, 2021, 11(5): 53–58, 64.
- [24] JIANG W, XU Y H, CHEN Z, et al. Fault diagnosis for rolling bearing using a hybrid hierarchical method based on scale-variable dispersion entropy and parametric t-SNE algorithm[J]. Measurement, 2022, 191: 110843.