



## 基于卷积长短时记忆网络的人体行为识别研究

孙彦玺<sup>1</sup>, 赵婉婉<sup>1</sup>, 武东辉<sup>1</sup>, 陈继斌<sup>1</sup>, 仇森<sup>2</sup>

(1. 郑州轻工业大学 建筑环境工程学院, 郑州 450002; 2. 大连理工大学 控制科学与工程学院, 辽宁 大连 116024)

**摘要:** 人体行为识别利用深度学习网络模型自动提取数据的深层特征, 但传统机器学习算法存在依赖手工特征提取、模型泛化能力差等问题。提出基于空时特征融合的深度学习模型(CLT-net)用于人体行为识别。采用卷积神经网络(CNN)自动提取人体行为数据的深层次隐含特征, 利用长短时记忆(LSTM)网络构建时间序列模型, 学习人体行为特征在时间序列上的长期依赖关系。在此基础上, 通过softmax分类器实现对不同人体行为分类。在DaLiAc数据集的实验结果表明, 相比CNN、LSTM、BP模型, CLT-net模型对13种人体行为的总体识别率达到了97.6%, 具有较优的人体行为识别分类性能。

**关键词:** 人体行为识别; 深度学习; 卷积神经网络; 长短时记忆网络; 模式识别; 可穿戴传感器

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 孙彦玺, 赵婉婉, 武东辉, 等. 基于卷积长短时记忆网络的人体行为识别研究[J]. 计算机工程, 2021, 47(10): 260-268.

**英文引用格式:** SUN Y X, ZHAO W W, WU D H, et al. Research of human activity recognition based on convolutional long short-term memory network[J]. Computer Engineering, 2021, 47(10): 260-268.

## Research of Human Activity Recognition Based on Convolutional Long Short-Term Memory Network

SUN Yanxi<sup>1</sup>, ZHAO Wanwan<sup>1</sup>, WU Donghui<sup>1</sup>, CHEN Jibin<sup>1</sup>, QIU Sen<sup>2</sup>

(1. College of Building Environment Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China;

2. School of Control Science and Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China)

**[Abstract]** Human activity recognition is a deep learning-based technology, which uses deep learning network models to automatically extract deep features of data. The traditional machine learning algorithms rely heavily on manual intervention during feature extraction, and exhibit a poor generalization ability. To address the problem, a deep learning model, CLT-net, is proposed based on space-time feature fusion for human activity recognition. CLT-net employs Convolution Neural Network (CNN) to extract the deep hidden features of human activity data automatically. Also, Long Short-Term Memory (LSTM) network is used to construct the time series model to learn the long-term dependence of human activity features on the time series. Finally, the softmax classifier is used to classify different human activities. The experimental results based on the public dataset, DaLiAc, show that CLT-net achieves an accuracy of 97.6% in the recognition of 13 kinds of human activities, outperforming the traditional models based on CNN, LSTM and BP. CLT-net has better classification performance of human activity recognition.

**[Key words]** human activity recognition; Deep Learning (DL); Convolutional Neural Network (CNN); Long Short-Term Memory (LSTM) network; pattern recognition; wearable sensors

**DOI:** 10.19678/j.issn.1000-3428.0060938

### 0 概述

人们的日常活动是构成社会生产、生活的重要组

成部分, 人体行为识别在日常生活中起着重要的作用,

被广泛应用在医疗康复、智能看护、运动监测、人机交

**基金项目:** 国家自然科学基金青年科学基金项目(61803072); 河南省科技攻关项目(182102210622); 河南省高等学校重点科研项目(19A413013); 郑州轻工业大学青年骨干项目(13501050002); 郑州轻工业大学博士科研项目(13501050009)。

**作者简介:** 孙彦玺(1996—), 男, 硕士研究生, 主研方向为深度学习、人体行为识别; 赵婉婉, 助理实验师; 武东辉, 讲师、博士; 陈继斌, 教授、硕士; 仇森, 讲师、博士。

**收稿日期:** 2021-02-25 **修回日期:** 2021-04-27 **E-mail:** w\_donghui@163.com

互等领域<sup>[1-2]</sup>。根据数据来源不同,人体行为识别分为基于视频图像的人体行为识别和基于可穿戴传感器的人体行为识别<sup>[3]</sup>。基于视频图像的人体行为识别是利用图像、视频处理等相关技术,通过对摄像设备获取的人体运动图像集合或视频片段进行分析,从而实现对人体行为识别和特定目标检测<sup>[4-5]</sup>。基于可穿戴传感器的人体行为识别能够从低级别的原始传感器数据中寻求关于人体行为中所蕴含的深层次知识,传感器数据主要由加速度计、陀螺仪、磁力计等传感器采集得到<sup>[6]</sup>。随着微电子机械系统和无线通信技术的飞速发展,这些传感器可以被集成在可穿戴设备、智能手机或智能手表中,极大地方便了人们日常携带,且在实际使用时不受场所和周围环境的限制,也不会给用户带来侵犯个人隐私的威胁<sup>[7]</sup>,促进了基于可穿戴传感器的人体行为识别在人们日常生活中的应用。

人体行为识别属于典型的模式识别问题,传统模式识别主要采用人工神经网络、支持向量机、决策树、朴素贝叶斯、K最近邻、隐马尔科夫模型等机器学习算法<sup>[8-9]</sup>。过去十几年,这些机器学习算法在人体行为识别的问题上取得了巨大的进步,但也存在一些不可回避的缺点。例如,利用传统机器学习算法进行人体行为识别时,需要提前手工提取人体行为数据特征,而手工特征提取受特定领域知识和人们已有知识与经验的限制<sup>[10]</sup>。一些浅层次特征(如均值、方差、频率等统计信息)只能用于识别人体行为的低级活动(如站立、行走、跑步等),很难识别出更为复杂和高级的人体行为(如洗盘子、使用吸尘器等),尤其是在当前多模态和高维传感器数据涌现的情况下,这些特征无法有效处理复杂的活动和实现对人体活动的准确分类<sup>[11]</sup>。

为进一步提高人体行为识别分类性能,同时又不依赖手工特征提取,本文提出基于空时特征融合的深度学习模型(CLT-net)的人体行为识别方法。CLT-net模型通过卷积神经网络(Convolutional Neural Network, CNN)自动提取数据特征,采用长短时记忆(Long Short-Term Memory, LSTM)网络学习时序数据的相关性优点并利用softmax分类器实现人体行为分类。

## 1 相关工作

近年来,深度学习技术得到蓬勃发展,在图像识别<sup>[12]</sup>、目标检测<sup>[13]</sup>、视频动作识别<sup>[14-15]</sup>、自然语言处理<sup>[16]</sup>、

时间序列预测<sup>[17]</sup>等领域取得了良好的效果。与传统机器学习算法不同,深度学习网络在很大程度上减轻了研究人员手工提取特征的工作量,模型通过更深层次的网络训练,可以自动提取到更高级别、更有意义的数据特征,使其在面对复杂的人体行为识别时更显优势<sup>[18]</sup>。

CNN和LSTM是现阶段最常用于人体行为识别的两种深度学习网络。文献[19]提出一种基于一维CNN方法来进行人体行为识别,该方法对基于智能手机加速度传感器走路、跑步、静止3种人体行为的识别准确率达到92.71%。文献[20]提出一种基于二维和三维的CNN模型学习人体行为数据特征,并结合手工提取特征用于人体行为识别,在UCI数据集上的实验结果表明,该方法的准确率达到96.95%。文献[21]设计一种长短时记忆递归神经网络(LSTM-RNN)来进行人体行为识别,该方法实现特征自动提取和融合,分类效果较好。文献[22]提出一种基于双向长短期记忆(BiLSTM)循环神经网络模型,并采用主成分分析(Principal Component Analysis, PCA)来降低数据集的维度,在mHealth数据集上的实验结果表明,PCA-BiLSTM模型的准确率达到97.64%。

## 2 人体行为识别模型

### 2.1 CNN模型

虽然深度学习网络在图像分类、人脸识别、自然语言处理等领域已经取得较好的效果,但是在序列信号分类应用上还没有得到一种公认的结构。因此,本文在LeNet-5基础上设计了CNN模型用于人体行为识别。相比LeNet-5,本文CNN模型输入层数据格式为24×410序列,并且在每个最大池化层之后加入批归一化层(Batch Normalization)、激活层(Leaky Relu)。CNN结构如图1所示,主要包括序列输入层、折叠层、CNN特征提取层(Convolution Maxpooling Batch Normalization Leaky Relu)、解折叠层、flatten层、全连接层、softmax分类层。从图1可以看出,CNN模型共包含3个CNN特征提取层。每个特征提取层的卷积层可以提取人体行为数据特征,是CNN模型的关键一层;最大池化层具有压缩数据、降低维度的作用;批归一化层对提取出的特征进行归一化处理;非线性激活层(Leaky Relu)可以促进批归一化之后特征的映射。全连接层可以减少提取特征的信息损失。Softmax分类层最终实现对人体行为的分类。

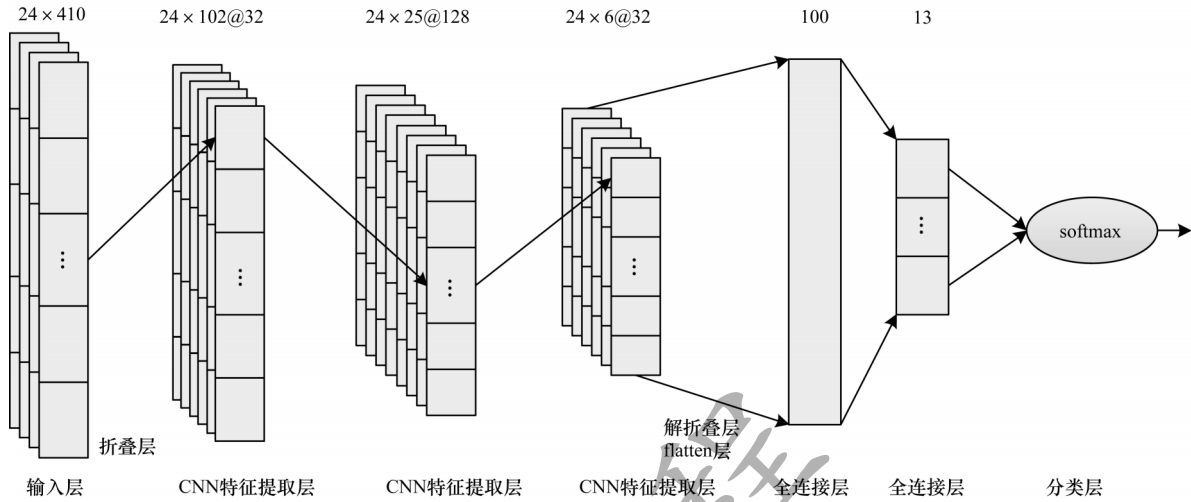


图1 CNN模型结构

Fig.1 Structure of CNN model

谷歌于2015年提出批归一化技术,其应用于深度神经网络训练不仅可以加快模型的收敛速度,而且一定程度缓解了深层网络中“梯度弥散”的问题,使训练的深度学习模型更稳定。Leaky Relu激活函数是为解决Relu输入值为负时,输出始终为0,同时一阶导数也始终为0,导致神经元参数不更新,神经元不学习的情况下出现的,定义如式(1)所示:

$$f(x) = \begin{cases} x, & x \geq 0 \\ s \times x, & x < 0 \end{cases} \quad (1)$$

其中: $s$ 为一个不小于1的非负数,当 $s$ 取0时,Leaky Relu激活函数退化为Relu函数。softmax分类层如式(2)所示:

$$s(x_i) = \frac{e^{x_i}}{\sum_{j=1}^K e^{x_j}}, i = 1, 2, \dots, K \quad (2)$$

其中: $x_i$ 为提取得到的人体行为数据的特征序列; $K$ 为人体行为类别数。softmax函数的分类结果代表输入样本被划分为每个类别时的所属概率,且所属概率和为1。

由于CNN能自动提取人体行为数据的深层特征,可以避免手工特征提取带来的诸多问题,因此本文将CNN特征提取层作为所提CLT-net网络模型的特征提取单元引入。CNN特征提取层实现特征提取的具体过程是3个卷积层的卷积核个数依次设为32、128、32,卷积核大小分别设为(1,11),(1,9),(1,7),步长为(1,2),即垂直方向步长为1,水平方向步长为2,同时模型采用same方式进行“padding”;3个最大池化层的池化核大小均为(1,3),步长为(1,2)。当单个样本的输入数据规模为 $24 \times 410 \times 1$ 时,经过3个CNN特征提取层得到的人体行为特征序列规模分别为 $24 \times 102 @ 32$ 、 $24 \times 25 @ 128$ 、 $24 \times 6 @ 32$ 。

## 2.2 LSTM模型

LSTM是循环神经网络(Recurrent Neural Network, RNN)的改进,由HOCHREITER和SCHMIDHUBER于1997年提出<sup>[23]</sup>。LSTM网络的核心部分是序列输入层和LSTM层,序列输入层可以将序列或时间序列数据输入网络,LSTM层可以学习序列数据时间步长之间的长期依赖关系,很好地解决RNN梯度消失问题<sup>[24]</sup>。由于LSTM是一种强大的时序信号处理和预测方法,人体行为传感器数据又属于时间序列上的信号,因此本文将LSTM层作为所提CLT-net网络模型的特征筛选单元引入。LSTM模型结构如图2所示,主要包括序列输入层、flatten层、LSTM层、全连接层、softmax分类层。

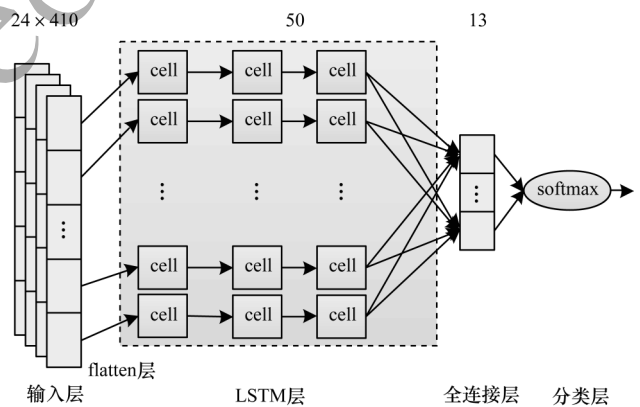


图2 LSTM模型结构

Fig.2 Structure of LSTM model

从图2可以看出,序列输入层样本大小为 $24 \times 410 \times 1$ ,经过flatten层将多维数据一维化后作为LSTM层输入,LSTM层隐藏单元数量设置为50,全连接层隐藏节点设置为13,最终由softmax分类层实现对不同人体行为分类。



LSTM层的细胞(cell)为输入数据提供时间依赖性,赋予了数据时间特征,LSTM网络通过细胞实现长期控制,进而用于时序信号的分类预测。细胞功能主要是通过遗忘门、输入门和输出门实现。LSTM层细胞内部结构如图3所示。

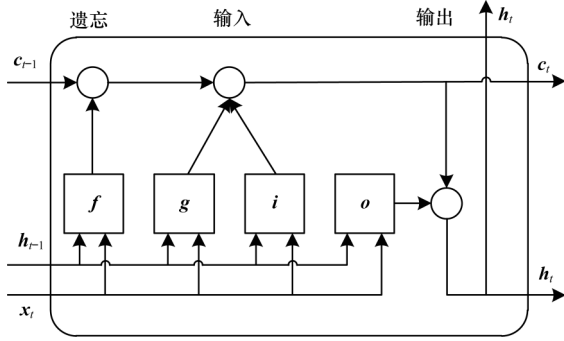


图3 LSTM细胞内部结构

Fig.3 The internal structure of LSTM cell

LSTM层可学习权值为输入权重 $W$ 、递归权重 $R$ 和偏差 $b$ 。矩阵 $W$ 、 $R$ 和 $b$ 分别是输入权重、递归权重和每个分量偏差的串联,如式(3)所示:

$$W = \begin{bmatrix} W_f \\ W_g \\ W_i \\ W_o \end{bmatrix}, R = \begin{bmatrix} R_f \\ R_g \\ R_i \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_f \\ b_g \\ b_i \\ b_o \end{bmatrix} \quad (3)$$

$t$ 时刻细胞状态输出和隐藏状态输出如式(4)、式(5)所示:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (4)$$

$$h_t = o_t \odot \sigma_c(c_t) \quad (5)$$

其中: $\odot$ 为Hadamard乘积(向量的元素相乘); $\sigma_c$ 为双曲正切函数(tanh)状态激活函数。

图3中 $t$ 时刻遗忘激活 $f_t$ 、输入激活 $i_t$ 、输出激活

$o_t$ 、候选单元输入 $g_t$ 如式(6)~式(9)所示:

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \quad (6)$$

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \quad (7)$$

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \quad (8)$$

$$g_t = \sigma_c(W_g x_t + R_g h_{t-1} + b_g) \quad (9)$$

其中: $h_{t-1}$ 为上一时刻隐藏状态的输出信息; $x_t$ 为当前时刻的输入信息。将 $h_{t-1}$ 与 $x_t$ 共同作为当前时间步的输入信息参与网络训练。这些信息经过门激活函数 $\sigma_g$ 后,最终得到输出介于 $[0,1]$ 的值。

遗忘激活 $f_t$ 越大,代表遗忘上一时刻细胞状态输出 $c_{t-1}$ 越少;输入激活 $i_t$ 越大,代表候选输入 $g_t$ 被写进当前时刻的信息越多,即遗忘激活 $f_t$ 和输入激活 $i_t$ 共同决定了当前时刻细胞状态输出 $c_t$ 对不同输入信息的接收程度;输出激活 $o_t$ 则确定了当前时刻隐藏状态的输出 $h_t$ ,以上控制策略便实现了人体行为数据在时间步长序列上的长期依赖性。

### 2.3 CLT-net人体行为识别模型

由于惯性传感器采集的人体行为数据可以看作是时间序列信号,以及当前机器学习算法严重依赖手工设计的特征,可能会导致信息利用不充分,无法有效实现对复杂人体活动识别等问题。本文提出了基于空时特征融合的深度学习模型(CLT-Net)用于人体行为识别。CLT-Net模型充分结合CNN自动提取数据深层特征和LSTM学习时序数据之间相关性的优点,采用与上述CNN模型相同的网络结构,同时使用LSTM层替换CNN模型的第一个全连接层。CLT-Net网络模型各项参数设置和规范函数选择与对应结构的CNN模块和LSTM模块一致,其结构如图4所示,主要包括序列输入层、折叠层、CNN特征提取层(Convolution Maxpooling Batch Normalization Leaky Relu)、解折叠层、flatten层、LSTM层、全连接层、softmax分类层。

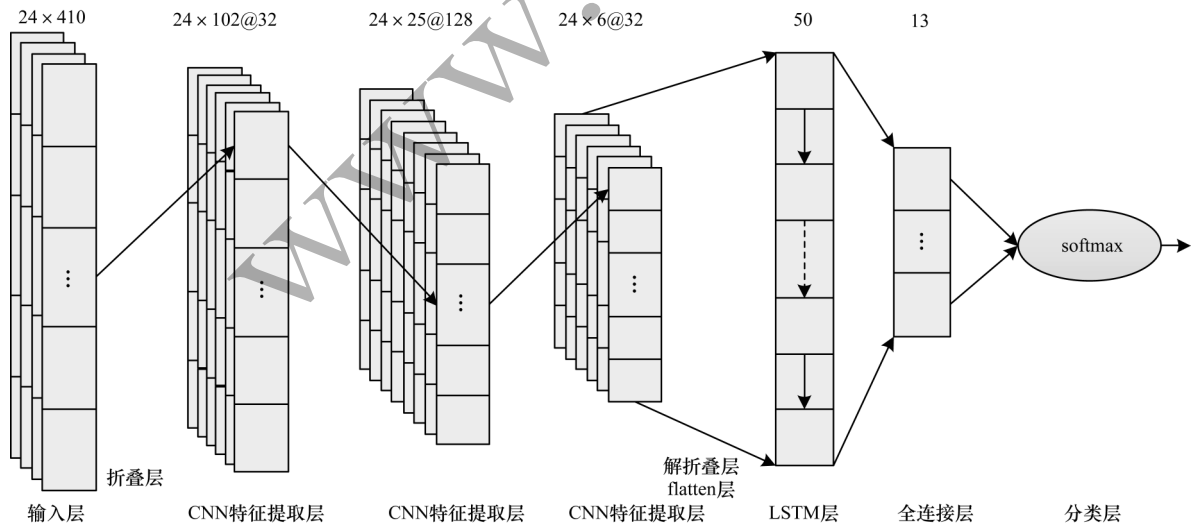


图4 CLT-net模型结构

Fig.4 Structure of CLT-net model

基于CLT-net模型的人体行为识别方法分类过程为:首先将输入的人体行为数据序列经过CNN模块进行二维空间上的特征提取,将得到的二维数据特征经过flatten层展开成一维后输入LSTM层进行时间序列上的特征筛选;然后再经过全连接层将筛选后得到的人体行为特征通过权重矩阵映射至样本标记空间;最终通过softmax层进行分类计算,并选择具有最大预测概率的类别作为输入数据样本的预测类别。模型在训练时会根据前向传播得到的预测类别与真实样本标签之间的误差,并依据模型采用的损失函数和优化器进行反向传播,以不断修正网络中的权值和偏置项,最终实现模型训练并得到较优模型。

### 3 实验

#### 3.1 实验数据集

本文基于可穿戴传感器数据进行人体行为识别研究,实验采用公开的DaLiAc (Daily Life Activities)数据集<sup>[25]</sup>。该数据集的数据采集过程由分别放置在受试者右臀部、胸部、右手腕和左脚踝4个6轴惯性传感器节点完成,每个传感器节点由一个三轴加速度计和一个三轴陀螺仪组成。加速度计的量程为 $\pm 6\text{ g}$ ,手腕、胸部、臀部传感器节点的陀螺仪范围为 $\pm 500(^{\circ})/\text{s}$ ,踝关节传感器节点的陀螺仪范围为 $\pm 2\,000(^{\circ})/\text{s}$ ,数据的采样频率为204.8 Hz。数据采集实验共有19名健康受试者参与(女性8名,男性11名,年龄 $26\pm 8$ 岁,身高 $177\pm 11\text{ cm}$ ,体重 $75.2\pm 14.2\text{ kg}$ ,偏差 $\text{mean} \pm \text{std}$ ),共采集了13项活动。活动及对应标签如表1所示。

表1 活动及对应标签

Table 1 Activities and corresponding labels

活动描述	标签
端坐	1
平躺	2
站立	3
洗盘子	4
吸尘	5
扫地	6
步行	7
上楼	8
下楼	9
跑步机跑步	10
测功单车骑车(50 W)	11
测功单车骑车(100 W)	12
跳绳	13

#### 3.2 实验运行环境

本文所有模型均在配置为Core i5-6500U CPU @ 3.20 GHz, 16 GB内存的计算机上进行训练和测试,该计算机系统为Windows 10专业版64位,并且所有模型使用Matlab2020b Deep Learning Toolbox 框架实现。

#### 3.3 实验参数

首先对人体行为数据进行样本划分,取滑动窗口长度为410(2倍的采样频率取整),且相邻窗口之间存在50%的数据重叠,因此单个样本序列大小为 $24\times 410$ (4个6轴传感器)。将经过数据分割后的样本打乱顺序,取前90%的样本作为训练集,剩下10%作为测试集。实验参数设置如表2所示,仿真实验时所有模型的初始化参数都使用相同配置,目的是使所有模型都在相对公平的条件进行对比,从而更利于准确反映CNN模型、LSTM模型和CLT-net模型的真实性能。

表2 实验参数设置

Table 2 Experimental parameters setting

参数	设置
CNN层权系数初始化	Kaiming法
LSTM层权系数初始化	正交法
全连接层权系数初始化	Kaiming法
优化器	Adam优化器
损失函数	交叉熵
初始学习率	0.001
样本序列大小	$24\times 410$
训练集样本数	20 088
测试集样本数	2 232
训练轮数	20
批尺寸	500
Leaky Relu因子	0.1

CNN层和全连接层的权系数初始化采用Kaiming法,该方法有利于加快模型的收敛速度,采用正交方法作为LSTM层的权系数初始化方法。所有模型的优化器都使用自适应矩估计优化算法(Adam算法)<sup>[26]</sup>,该方法具有更快的收敛速度和更低的内存消耗需求,并且在模型训练过程中可以不使用验证集。

#### 3.4 实验结果与分析

##### 3.4.1 模型收敛速度

LSTM模型、CNN模型和CLT-net模型是在训练集上进行训练的,模型训练过程的准确率对比如图5所示。从图5可以看出,3种模型训练过程的准确率

**Fig.8 Prediction classification confusion matrix of CNN model on test set**





从3个模型的t-SNE可视化图形中可以看出,基于CLT-net模型的t-SNE数据可视化将13种人体行为很好分类(13种人体行为分别被聚成一簇),说明CLT-net模型具有优越的人体行为分类识别性能。CNN模型和LSTM模型次之。

#### 3.4.4 模型评价指标

为更好说明CLT-net模型的泛化能力,进一步统计了LSTM、CNN、CLT-net和传统BP模型在测试集上测试结果的宏查准率(macro precision)、宏查全率(macro recall)和宏F1值(macro F1-score),即对13种人体行为类别中每类行为的精确率、召回率和F1-score进行求和再取平均值<sup>[29]</sup>,4种模型的评价指标对比如表3所示。

表3 BP、LSTM、CNN、CLT-net模型的评价指标对比  
Table 3 Evaluation indexes comparison between BP, LSTM, CNN, CLT-net models

模型	宏查准率	宏查全率	宏F1值
BP	0.533 2	0.523 8	0.510 3
LSTM	0.769 6	0.722 8	0.737 1
CNN	0.962 1	0.961 4	0.961 2
CLT-net	0.974 6	0.971 6	0.972 7

精确率评价指标是衡量正确预测出的正样本数占实际预测出正样本数的比例;召回率评价指标是衡量正确预测出的正样本数占总正样本数的比例;F1-score评价指标作为精确率和召回率指标的调和平均值。从表3可以看出,CLT-net模型的精确率、召回率和F1-score评价指标均最高,而BP模型的所有指标均最低,表明CLT-net模型对不同人体行为的识别分类结果具有良好的稳定性与可靠性。

## 4 结束语

本文提出基于空时特征融合技术的深度学习模型CLT-net。该模型结合CNN可以自动提取数据特征和LSTM能够学习时序数据之间的相关性优点,并利用softmax分类器对人体行为进行分类。在DaLiAc数据集上的实验结果表明,相比LSTM、CNN、BP模型,CLT-net模型收敛速度更快且人体行为识别分类性能更优。后续将通过构建轻量级的深度学习模型,对基于传感器的人体行为识别方法进行优化,进一步提高特征识别准确率。

### 参考文献

- [1] FU B, DAMER N, KIRCHBUCHNER F, et al. Sensing technology for human activity recognition: a comprehensive survey[J]. IEEE Access, 2020, 8: 83791-83820.
- [2] HASSAN M M, UDDIN M Z, MOHAMED A, et al. A robust human activity recognition system using smartphone sensors and deep learning[J]. Future Generation Computer Systems, 2018, 81: 307-313.
- [3] DANG L M, MIN K, WANG H, et al. Sensor-based and vision-based human activity recognition: a comprehensive survey[J]. Pattern Recognition, 2020, 108: 107561.
- [4] TU Z, XIE W, QIN Q, et al. Multi-stream CNN: learning representations based on human-related regions for action recognition[J]. Pattern Recognition, 2018, 79: 32-43.
- [5] LI J, XIE X, PAN Q, et al. SGM-Net: skeleton-guided multimodal network for action recognition[J]. Pattern Recognition, 2020, 104: 107356.
- [6] BULLING A, BLANKE U, SCHIELE B. A tutorial on human activity recognition using body-worn inertial sensors[J]. ACM Computing Surveys, 2014, 46(3): 1-33.
- [7] FERHAT A, SAMER M, MARIAM D, et al. Physical human activity recognition using wearable sensors[J]. Sensors, 2015, 15(12): 31314-31338.
- [8] LIMA W S, SOUTO E, EL-KHATIB K, et al. Human activity recognition using inertial sensors in a smartphone: an overview[J]. Sensors, 2019, 19(14): 235-244.
- [9] JOBANPUTRA C, BAVISHI J, DOSHI N. Human activity recognition: a survey[J]. Procedia Computer Science, 2019, 155: 698-703.
- [10] WANG J, CHEN Y, HAO S, et al. Deep learning for sensor-based activity recognition: a survey[J]. Pattern Recognition Letters, 2017, 119: 3-11.
- [11] NWEKE H F, TEH Y W, AL-GARADI M A, et al. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: state of the art and research challenges[J]. Expert Systems with Applications, 2018, 105: 233-261.
- [12] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2021-01-20]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [13] 寇大磊, 权冀川, 张仲伟. 基于深度学习的目标检测框架进展研究[J]. 计算机工程与应用, 2019, 55(11): 25-34.
- [14] KOU D L, QUAN J C, ZHANG Z W. Research on progress of object detection framework based on deep learning[J]. Computer Engineering and Applications, 2019, 55(11): 25-34. (in Chinese)
- [15] TU Z, LI H, ZHANG D, et al. Action-stage emphasized spatio-temporal VLAD for video action recognition[J]. IEEE Transactions on Image Processing, 2019, 28(6): 2799-2812.
- [16] YONG D, YUN F, LIANG W. Representation learning of temporal dynamics for skeleton-based action recognition[J]. IEEE Transactions on Image Processing, 2016, 25(7): 3010-3022.
- [17] YOUNG T, HAZARIKA D, PORIA S, et al. Recent trends in deep learning based natural language processing[J]. IEEE Computational Intelligence Magazine, 2018, 13(3): 55-75.
- [18] LI X, YI X, LIU Z, et al. Application of novel hybrid deep learning model for cleaner production in a paper industrial wastewater treatment system[J]. Journal of Cleaner Production, 2021, 294: 1-12.
- [19] ABDULMAJID M, JAE-YOUNG P. Deep recurrent neural networks for human activity recognition[J]. Sensors, 2017,



- 17(11):2556.
- [19] LEE S M, YOON S M, CHO H. Human activity recognition from accelerometer data using convolutional neural network [C]//Proceedings of 2017 IEEE International Conference on Big Data and Smart Computing. Washington D. C. , USA: IEEE Press, 2017: 131-134.
- [20] HA Q D, TRAN M T. Activity recognition from inertial sensors with convolutional neural networks [C]//Proceedings of International Conference on Future Data and Security Engineering. Berlin, Germany: Springer, 2017: 285-298.
- [21] 范长军, 高飞. 基于可穿戴传感器的普适化人体活动识别[J]. 传感技术学报, 2018, 31(7): 1124-1131.  
FAN C J, GAO F. Human daily activity recognition based on wearable sensors[J]. Chinese Journal of Sensors and Actuators, 2018, 31(7): 1124-1131. (in Chinese)
- [22] ALJARRAH A A, ALI A H. Human activity recognition using PCA and BiLSTM recurrent neural networks [C]//Proceedings of the 2nd International Conference on Engineering Technology and its Applications. Washington D. C. , USA: IEEE Press, 2019: 156-160.
- [23] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [24] YU Y, SI X, HU C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019, 31(7): 1235-1270.
- [25] LEUTHEUSER H, SCHULDHAUS D, ESKOFIER B M. Hierarchical, multi-sensor based classification of daily life activities; comparison with state-of-the-art algorithms using a benchmark dataset[J]. PLoS One, 2013, 8(10): 75196-75207.
- [26] 赵小强, 宋昭漾. Adam优化的CNN超分辨率重建[J]. 计算机科学与探索, 2019, 13(5): 858-865.  
ZHAO X Q, SONG Z Y. Adam optimized CNN super-resolution reconstruction [J]. Journal of Frontiers of Computer Science and Technology, 2019, 13(5): 858-865. (in Chinese)
- [27] 魏世超, 李歆, 张宜弛, 等. 基于E-t-SNE的混合属性数据降维可视化方法[J]. 计算机工程与应用, 2020, 56(6): 66-72.  
WEI S C, LI X, ZHANG Y C, et al. Dimension reduction and visualization of mixed-type data based on E-t-SNE [J]. Computer Engineering and Applications, 2020, 56(6): 66-72. (in Chinese)
- [28] 马吉, 刘瑞, 张建霞. 基于改进t-SNE算法的人体运动数据关键帧提取[J]. 计算机工程, 2016, 42(5): 258-262.  
MA J, LIU R, ZHANG J X. Key frame extraction for human motion data based on improved t-SNE algorithm[J]. Computer Engineering, 2016, 42(5): 258-262. (in Chinese)
- [29] 陈明豪, 祝跃飞, 芦斌, 等. 基于Attention-CNN的加密流量应用类型识别[J]. 计算机科学, 2021, 48(4): 325-332.  
CHEN M H, ZHU Y F, LU B, et al. Classification of application type of encrypted traffic based on attention-CNN [J]. Computer Science, 2021, 48(4): 325-332. (in Chinese)

编辑 薛晋栋