

## 基于热释电红外传感器的人体动作识别方法

徐晓冰, 左涛涛\*, 孙百顺, 李奇越, 吴 刚

(合肥工业大学 电气与自动化工程学院, 安徽 合肥 230009)

**摘要:** 针对目前人体动作识别技术中存在的隐私暴露、技术复杂度高和识别精度低等相关问题, 提出了一种基于热释电红外 (PIR) 传感器的人体动作识别方法。首先, 采用一组安置在天花板上经过视场调制的 PIR 传感器采集人体运动时散发的红外热辐射信号, 将传感器输出的电压模拟信号进行滤波放大后通过 ZigBee 无线模块传送到 PC 端打包成原始数据集; 其次, 将原始数据的两路传感器输出数据进行特征融合, 对融合后的数据做标准化处理封装为训练集和测试集; 然后, 基于数据的特征提出一种两层级联的混合深度学习网络模型作为人体动作的分类算法, 第一层采用一维卷积神经网络 (1DCNN) 对数据进行特征提取, 第二层采用门控循环单元 (GRU) 保存历史输入信息防止丢失有效特征; 最后, 利用训练集来训练该网络模型得出参数最优的分类模型, 通过测试集验证模型的正确性。实验结果表明, 提出的该动作识别技术模型对基本动作分类的准确率高于 98%, 与图像动作识别或穿戴式设备动作识别相比, 实现了实时、便捷、低成本和高保密性的高精度人体动作识别。

**关键词:** 热释电红外传感器; 动作识别; 一维卷积神经网络; 门控循环单元

**中图分类号:** TP212.9; TP391.4 **文献标志码:** A **DOI:** 10.3788/IRLA20210188

## Human motion recognition method based on pyroelectric infrared sensor

Xu Xiaobing, Zuo Taotao\*, Sun Baishun, Li Qiyue, Wu Gang

(School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China)

**Abstract:** Aiming at the privacy exposure, high technical complexity, and low recognition accuracy existing in the current human motion recognition technology, this paper proposed a human motion recognition method based on a pyroelectric infrared (PIR) sensor. Firstly, a set of PIR sensors placed on the ceiling and modulated by the field of view were used to collect the infrared heat radiation signal emitted by the human body when moving, and the voltage analog signal output by the sensor was filtered and amplified, and then transmitted to the PC through the ZigBee wireless module and packaged into raw data. Secondly, the two-way sensor output data of the original data feature was fused, and the fused data was standardized and packaged into training dataset and test dataset. Then, a two-layer cascaded hybrid deep learning network was proposed to be a classification algorithm of human motion based on the characteristics of the data. The first layer used one-dimensional convolutional neural network (1DCNN) to extract features from the data, and the second layer used gated recurrent unit (GRU) to save historical input information to prevent loss of valid features. Finally, the training dataset was used to train the network model to obtain a classification model with the best parameters, and the correctness of the model was

收稿日期: 2021-03-19; 修订日期: 2021-04-12

作者简介: 徐晓冰, 男, 副教授, 硕士生导师, 主要从事物联网控制系统、智能信息处理与智能系统方面的研究。

通讯作者: 左涛涛, 男, 硕士生, 主要从事智能信息处理与智能系统、物联网控制系统方面的研究。

verified through the test dataset. The experimental results show that the accuracy of the proposed motion recognition technology model for basic motion classification is higher than 98%. Compared with image motion recognition or wearable device motion recognition, it realizes high-precision human motion recognition with real-time, convenience, low cost and strong confidentiality.

**Key words:** pyroelectric infrared sensor; action recognition; one-dimensional convolutional neural network; gated recurrent unit

## 0 引言

近年来,人类动作识别因能在许多领域上应用而引起广泛的研究兴趣,精准的动作识别方法可以对人体行为进行判断,在运动分析、辅助医疗、虚拟现实和人机交互等领域得到广泛的应用。目前动作识别的方法主要可分为三类,分别利用图像、穿戴式设备和电磁波进行识别。基于计算机视觉的人体动作识别<sup>[1-3]</sup>是最热门的研究方向,Sandhya等<sup>[4]</sup>和裴晓敏<sup>[5-6]</sup>等通过人体骨架的几何特征来识别人体行为,但视觉传感器易引发私密场合的隐私泄露问题;一类是基于惯性传感器的人体动作识别<sup>[7-8]</sup>,利用加速度计和陀螺仪等<sup>[9]</sup>检测运动的速度和角度信息来提取姿态信息,但专用传感器种类和数量太多导致安装和携带都不方便;近年来有一种新的方法是基于电磁波信号的动作识别,如WiFi<sup>[10-11]</sup>和人体红外热辐射,使用WiFi信号识别复杂度较高且准确度较低,所以利用人体红外热辐射信号来进行动作识别是很好的选择。热释电红外(pyroelectric infrared, PIR)传感器作为一种无源传感器,它可以检测到人体运动时红外热辐射的变化,还具有低成本、低功耗和高灵敏度的特点,最重要的是可以保护隐私。

De<sup>[12]</sup>等提出了一种基于四 PIR 传感器的系统,结合改进的 MCP 算法,可以识别室内环境中八个方向的人体运动检测。Pourpanah<sup>[13]</sup>等人利用带有随机采样窗口掩膜的 PIR 传感器来感知人体的热变化,并提取信号时域统计特征,利用 QFAM-GA 方法识别人

体行走动作的异常与否。Sun<sup>[14]</sup>等提出了一种双模式感知系统,既可以检测运动目标,也可利用运动控制单元主动检测静止的人体对象,从而可以有效识别包含静止和运动目标的场景。Guan<sup>[15]</sup>等设计了一套增强型红外传感系统,针对侧视传感模块和俯视传感模块使用不同的可见光掩膜,可显著增强特征的敏感度,利用 GM-HMM 算法可以识别走、坐、跑等日常活动,平均准确率达到 91.6%。Yang<sup>[16]</sup>等提出了一种基于分布式 PIR 传感器和多维特征融合的跌倒检测系统,以人摔倒过程中的质心变化、速度变化、人的面积变化及相邻帧的方差变化作为特征值,最后利用 SVM 实现了跌倒动作的识别,平均准确率可以达到 93%。

考虑到 PIR 传感器输出的模拟信号本质上是一种时间序列数据,因此文中结合了一维卷积神经网络(one-dimensional convolutional neural network, 1DCNN)与循环神经网络(recurrent neural network, RNN)对动作进行分类,其中 RNN 选取网络结构较为简单的门控循环单元(gated recurrent unit, GRU)。该方法利用人体运动时 PIR 传感器输出的原始信号处理之后的数据作为 1DCNN-GRU 网络模型的输入,训练后能准确地识别出坐、站、走、跑和摔倒这五种不同的基本动作。

## 1 人体动作信号采集及分析

### 1.1 总体设计

提出的系统由传感器节点、MCU 数据采集单元和无线数据传输模块构成,图 1 为整体结构示意图。

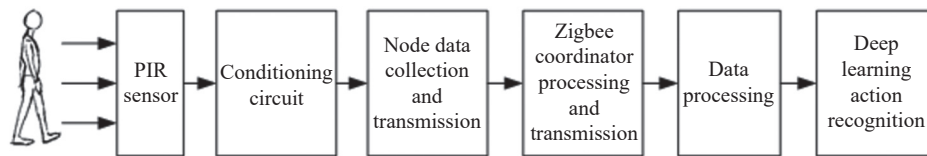


图 1 系统整体结构图

Fig.1 System overall structure diagram

传感器节点由两个使用幕帘型菲涅耳透镜的 PIR 传感器组成,负责感应人体运动的信息。MCU 数据采集单元包括信号调理电路模块和 ADC 采集模块,调理电路对 PIR 的输出信号进行滤波和放大,经 ADC 转换为数字信号由 ZigBee 无线传输网络实时传送到 PC 端,对数据进行预处理并打包封装好,作为深度学习网络模型的训练集和测试集。

### 1.2 传感器节点设计及信号特征分析

#### 1.2.1 PIR 传感器工作原理及节点设计

PIR 传感器通常由热释电元件和菲涅耳透镜等组成。人体体温一般在 37 °C 左右,会发出特定波长(8~14 μm)的红外线。热释电元件在接收到人体红外辐射时会电荷失去平衡向外释放电荷,即热释电效应。菲涅耳透镜的作用一是聚集人体红外能量投射到热释电元件表面,扩大检测距离;二是将视场分为可见区与暗区,保证投射到热释电元件表面上的红外辐射能量信号是变化的。当人体在 PIR 传感器感应区运动时,两片热释电元件接收的人体热辐射能量值不同,便会产生一个变化的电压输出信号,该信号包含人体及运动的相关特征。笔者设计的实验中,PIR 传感器采用二元被动热释电传感器 D203 S,其包含一对反向串联的热势电元件,可以消除环境温度波动的影响。热释电探头的内部结构如图 2 所示。

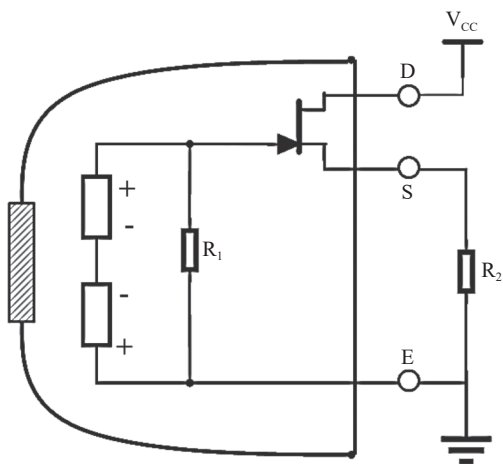


图 2 PIR 内部结构图

Fig.2 Internal structure of PIR

系统的传感器节点依据俯视视角安装,与侧视视角相比更能避免一些障碍物的遮挡,同时也能收集到

更多人体运动的信息。图 3、4 分别是传感器的节点设计模型和实验场景。将两个 PIR 传感器安放在距离地面 2.5 m 高的天花板上,PIR 传感器之间的距离设置为 2.5 m,并对视场角进行调制,最终使得检测区域面积为 2.5 m×0.5 m。

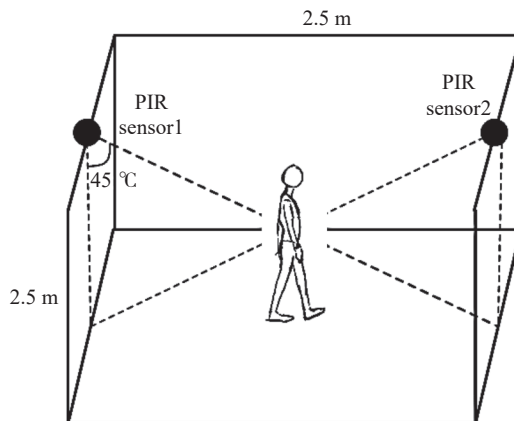


图 3 传感器节点参考模型

Fig.3 Sensor node reference model

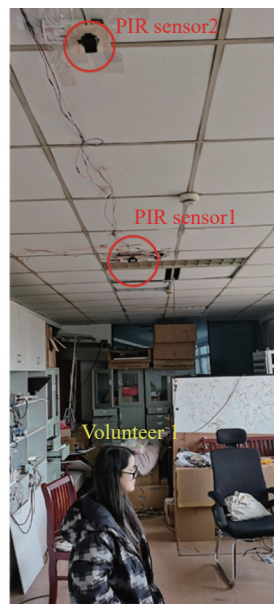


图 4 实验实景图

Fig.4 Experimental real picture

#### 1.2.2 信号特征分析

基于上述的热释电红外传感器工作原理,当人在 PIR 传感器监测范围内运动时,会导致 PIR 传感器接收到的热辐射发生变化,从而产生电压波形信号。经研究发现:在相同类型的 PIR 传感器及菲涅耳透镜

下,其捕捉到的人体运动信号具有很明显的三个特征。一是当人的动作经过由菲涅耳透镜划分而成的明暗区越多,其采集到的波形波峰的个数越多;二是当人的动作经过视场明暗区的速度越快,其采集到的波形的斜率的绝对值就越大。三是因笔者采用的 PIR 传感器是二元型,采集的信号具有方向性,可以判断动作在检测区域内的运动方向。由此,可以通过采集到的电压模拟信号的时域及频域特征信息来识别其动作类型。

## 2 深度学习网络设计

### 2.1 1DCNN 和 GRU 网络

卷积神经网络 (convolutional neural network, CNN) 是由 Lecun 等<sup>[17]</sup>在深度学习研究中首次提出的一种深度学习架构。1DCNN 网络在训练过程中能够减少神经网络模型参数的数量,通过卷积核的权值共享使得运算量大幅度降低,而且不会因数据的维度压缩而导致大量时序特征信息损失。结构图如图 5 所示。

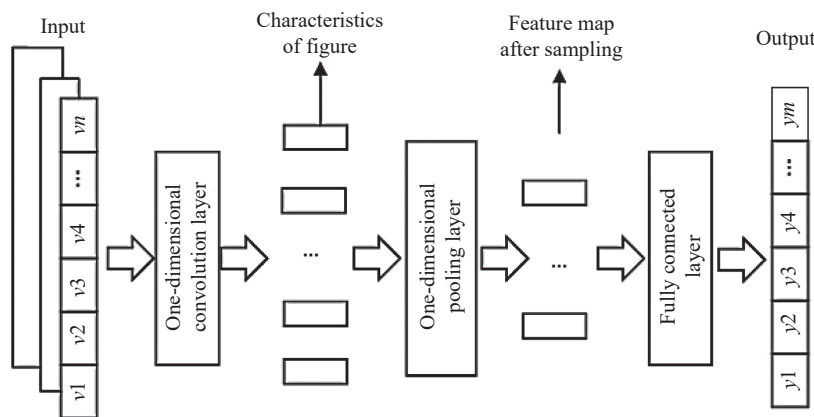


图 5 一维卷积神经网络结构图

Fig.5 Diagram of one-dimensional convolutional neural network structure

RNN 擅长处理时间序列数据,但在输入序列的维度很大时可能会出现梯度消失的情况。长短期记忆网络 (long short-term memory, LSTM) 和 GRU 对 RNN 进行优化,解决了梯度消失的问题。GRU 是 LSTM 的一种变体<sup>[18]</sup>,比 LSTM 少了一个门函数,参数的数量远少于 LSTM,因此 GRU 训练速度更快且成本更低。GRU 网络模型中有两个门:更新门和重置门。更新门控制当前时刻输出的状态中要保留多少历史状态以及保留多少当前时刻的候选状态;重置门决定当前时刻的候选状态是否需要依赖上一时刻的网络状态以及需要依赖多少<sup>[19]</sup>。具体结构如图 6 所示。

式中:  $\odot$  表示两个向量相连;  $*$  表示矩阵的乘积;  $r_t$  为重置门,  $z_t$  为更新门;  $W_r$ 、 $W_z$ 、 $W_h$  为权重参数;  $x_t$  为  $t$  时刻 GRU 的输入;  $\tilde{h}_t$  表示一个 GRU 单元在  $t$  时刻的激活值;  $h_t$  表示当前 GRU 单元的激活值;  $\sigma$  为 sigmoid 函数。

用公式可以表达为:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (1)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (2)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (3)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t * h_{t-1}, x_t]) \quad (4)$$

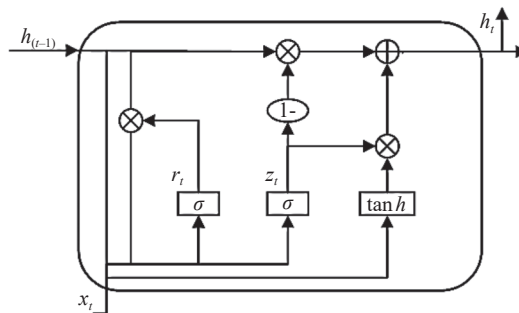


图 6 GRU 网络结构图

Fig.6 Diagram of GRU network structure

### 2.2 1DCNN-GRU 深度学习网络

文中将 1DCNN 网络与 GRU 网络各自的特点相

结合,采用了一种 1DCNN-GRU 的网络模型对人体动作进行分类识别。该设计中的网络模型输入参数是  $1 \times 1200$  维的特征向量,由于特征点较多,因此首先选择三层 1DCNN 神经网络进行深层特征提取,实现对特征量进行特征重构与特征向量降维,且同时保证不破坏特

征向量的时序性,然后利用 GRU 网络的非线性特征学习能力对卷积后的特征向量进行学习,得到可表述人体动作的特征集。最后通过全连接层、softmax 层对结果进行分类。网络结构如图 7 所示。

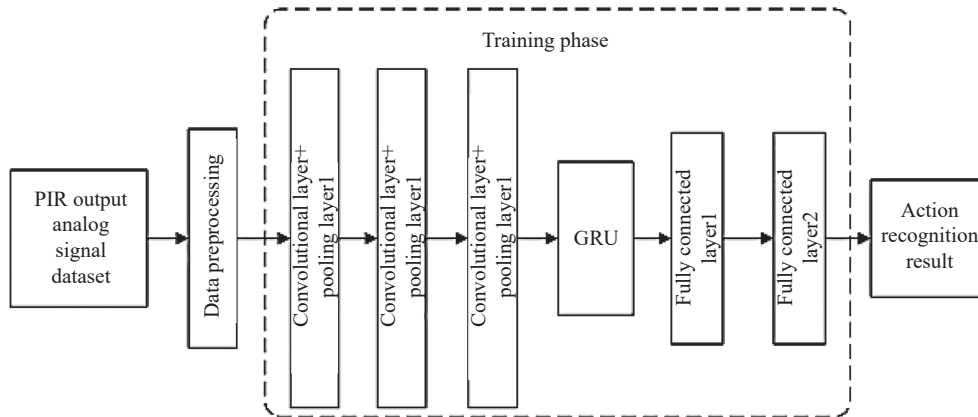


图 7 1DCNN-GRU 模型结构图

Fig.7 Structure diagram of 1DCNN-GRU model

### 3 实验及结果分析

文中实验提出的动作识别算法的流程图如图 8 所示,首先对一组 PIR 传感器采集的数据进行预处理

理,去除其他热源干扰,并将数据进行特征融合构建数据集。然后利用深度学习神经网络学习并建立了动作检测模型,最后使用机器学习分类算法来完成动作的识别。

#### 3.1 数据集

以实验采集志愿者在 PIR 传感器下做指定动作时的实时输出信号作为数据集。为保证实验的多样性,排除偶然性,招募五名志愿者(3 名男性,2 名女性)协助测量数据。要求志愿者在两个 PIR 传感器视场路径的中心区域完成指定动作,每位志愿者依次进行坐、站、走、跑和摔倒这五种动作。在检测现场无规则摆放各种障碍物,并且让每位志愿者以不同的速度、动作幅度和方向做上面五种动作。传感器节点采集一次数据的时间长度为 6 s,采样频率为 100 Hz,并通过计算机的 Matlab 软件来查看和保存实时数据。将两传感器的实时输出数据拼接形成数据集中的一个样本数据,该数据集中包含 1500 组样本数据,五种动作样本各 300 组。选取每种样本中的 200 组数据组成容量 1000 组的训练集,剩下的 500 组作为测试集来验证网络结构的有效性。图 9 为某一志愿者五种动作的原始数据。

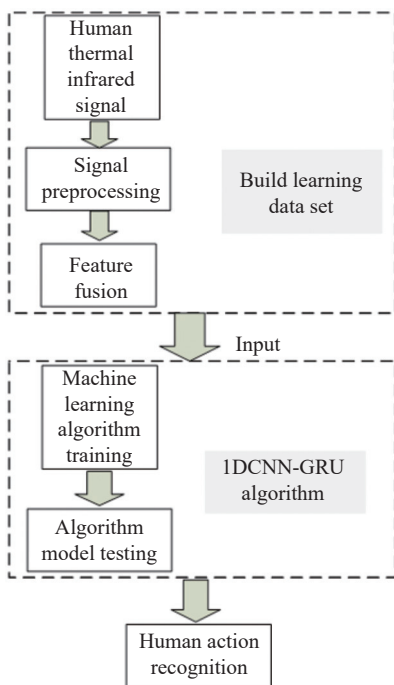


图 8 算法实现流程图

Fig.8 Flowchat of algorithm implementation

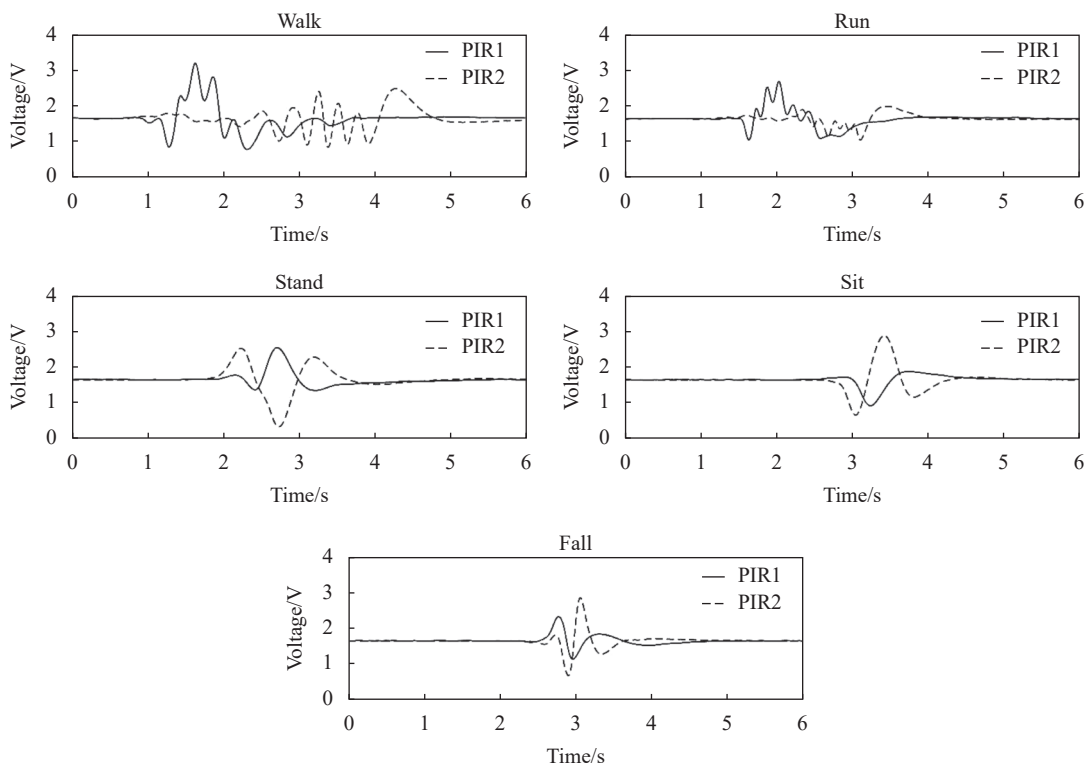


图 9 五种动作的原始信号图

Fig.9 Original signal diagram of five actions

### 3.2 数据预处理

由于五位志愿者行为习惯及体型都不同,导致样本数据的数值变化范围存在差异,需要对每个样本数据进行标准化处理。处理方法采用最大最小标准化,将样本数据映射到 (0,1) 区间之中,转换函数用公式表示为:

$$x'_t = \frac{x_t - \min(x_t)}{\max(x_t) - \min(x_t)} \quad (5)$$

式中:  $x_t$  表示某一样本数据;  $x'_t$  表示预处理后的样本数据;  $\min(x_t)$  表示一组样本数据的最小值;  $\max(x_t)$  表示一组样本数据的最大值。经过归一化之后再对数据进行网络训练可以加快梯度下降求最优解的速度,减少训练时间。

将经过归一化处理的两路信号进行特征融合,即将特征向量进行连接。若有  $n$  个输入特征向量  $x_1, x_2, \dots, x_n$ , 其维数分别为  $p_1, p_2, \dots, p_n$ , 则输出特征向量  $y$  的维数为  $\sum_{i=1}^n p_i$ 。最后将数据集加上标签,用训练集输入神经网络中进行训练,并用测试集测试模型训练效果。

### 3.3 实验结果及分析

实验中网络设计采用的是 Tensorflow+Keras 框

架,表 1 是具体网络结构设置。文字模型使用自适应性矩估计 (adaptive moment estimation, Adam) 优化函数更新模型的权重与偏置。为防止过拟合,设置 GRU 层的 dropout 值<sup>[20]</sup> 为 0.05。

表 1 1DCNN-GRU 模型参数设置

Tab.1 Parameter settings for 1DCNN-GRU model

Layer(type)	Output shape	Parameter
Conv1 d_1(Conv1D)	(None,1 200,256)	1 536
Max_pooling1d_1	(None,300,256)	0
Conv1 d_2(Conv1D)	(None,300,128)	98 432
Max_pooling1d_2	(None,75,128)	0
Conv1 d_3(Conv1D)	(None,75,64)	24 640
Max_pooling1d_3	(None,18,64)	0
gru_1(GRU)	(None,256)	246 528
dense_1(Dense)	(None,32)	8 224
dense_2(Dense)	(None,5)	165

神经网络学习训练的过程将参数 epochs 设置为 300,即训练 300 轮。图 10 是经过 300 次迭代之后得到的模型损失函数,图 11 是模型的训练准确度。训练结果显示:随着神经网络训练的进行,损失函数逐

渐减小,训练精度逐渐提高,未发生过拟合,并且在迭代 150 次左右损失值下降到零,训练精度达到百分之百。在变化过程中出现的波动表示模型并不是在向着同一个方向学习,因此会出现波动。但最终准确率和损失都趋于平稳,即模型收敛。

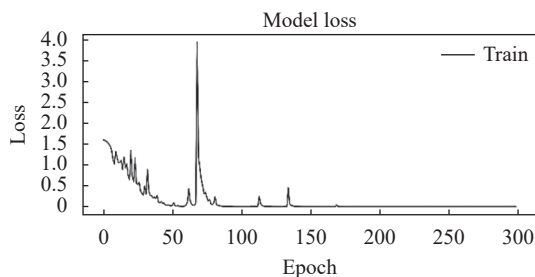


图 10 模型的损失函数

Fig.10 Loss function of the model

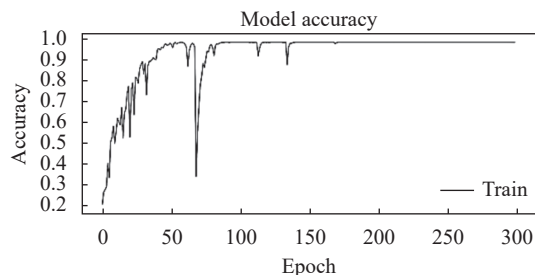


图 11 模型的训练准确度

Fig.11 Training accuracy of the model

表 2 是利用测试集在 1DCNN-GRU 模型下进行测试的准确率。实验结果显示:该模型的平均分类准确率达到 98.6%,达到了较好的识别效果。其中可以看出摔倒动作的识别率偏低,因为摔倒动作是一种复杂的全身动作,会导致 PIR 传感器采集热辐射信号变化的特征信息较多,因此识别率偏低。实验结果说明 1DCNN-GRU 网络模型对于文中的五种基础动作能够进行有效的区分。

在文中的研究中也对不同模型的性能进行了对比研究,表 3 所示为不同模型识别动作的训练时长及识别准确率的对比结果。GRU 网络与传统的 LSTM 网络相比较具有更好的优势,在保证识别准确率几乎相同的前提下,GRU 网络模型可以减少训练参数,加快收敛速度。尽管 1DCNN 的训练时长最短,但其识别的准确率还是较低的。因此在将特征向量输入到 GRU 网络之前,可利用卷积和池化进一步简化计算,在保证不丢失信号特征的同时得到性能最佳的训练

表 2 五种动作分类的准确率

Tab.2 Accuracy of 5 types of action classification

Action category	Recognition accuracy
Run	98%
Walk	100%
Sit	100%
Stand	100%
Fall	95%

表 3 不同模型的性能对比

Tab.3 Performance comparison of different models

Network model	Accuracy	Training time/s
1DCNN	93.8%	178
GRU	88.6%	467
LSTM	91.8%	543
1DCNN-LSTM	98.6%	245
1DCNN-GRU	98.8%	195

模型。最终实验结果表明:通过结合 1DCNN 和 RNN 网络之后的模型既减小了运算负荷,又更能捕获实验数据的时序特征。因此文中提出的方法达到了较好的动作识别效果。

## 4 结 论

文中针对分布式 PIR 传感器检测系统提出了一种基于时域卷积与循环神经网络的混合网络模型进行人体动作识别的算法,该算法可以有效识别短时活动。提出的 1DCNN-GRU 网络模型既可以通过卷积神经网络压缩数据维度加快运算速度,又能让 GRU 网络充分利用时间序列的相关性。通过与其他几种传统方法相比可以看出:文中使用的 1DCNN-GRU 模型效果更好,在保证识别准确率高的大大提升了性能。

此研究还存在一些不足,文中提出的方法只适用于室内环境对单个目标的基础动作的识别,多目标情况下采集到的信号复杂度较高,如何找到其中能准确表述各目标以及动作的特征信息将成为未来的研究方向。

## 参考文献:

[1] Patel C I, Labana D, Pandya S, et al. Histogram of oriented gradient-based fusion of features for human action recognition in

- action video sequences [J]. *Sensors*, 2020, 20(24): 7299.
- [2] Ma Shiwei, Liu Lina, Fu Qi, et al. Using PHOG fusion features and multi-class Adaboost classifier for human behavior recognition [J]. *Optics and Precision Engineering*, 2018, 26(11): 2827-2837. (in Chinese)
- [3] Li Qinghui, Li Aihua, Cui Zhigao, et al. Action recognition via restricted dense trajectories and spatio-temporal co-occurrence feature [J]. *Optics and Precision Engineering*, 2018, 26(1): 230-237. (in Chinese)
- [4] Sandhya R S, Apparao N G, Usha S V. Kinematic joint descriptor and depth motion descriptor with convolutional neural networks for human action recognition [J]. *Materials Today: Proceedings*, 2020, 37(2): 3164-3173.
- [5] Pei Xiaomin, Fan Huijie, Tang Yandong. Action recognition method of spatio-temporal feature fusion deep learning network [J]. *Infrared and Laser Engineering*, 2018, 47(2): 0203007. (in Chinese)
- [6] Pei Xiaomin, Fan Huijie, Tang Yandong. Two-person interaction recognition based on multi-stream spatio-temporal fusion network [J]. *Infrared and Laser Engineering*, 2020, 49(5): 20190552. (in Chinese)
- [7] Liu S Q, Zhang J C, Zhang Y Z, et al. A wearable motion capture device able to detect dynamic motion of human limbs [J]. *Nature Communications*, 2020, 11(1): 5615.
- [8] Su Benyue, Zheng Dandan, Tang Qingfeng, et al. Human daily short-time activity recognition method driven by single sensor data [J]. *Infrared and Laser Engineering*, 2019, 48(2): 0226003. (in Chinese)
- [9] Wang Zhenyu, Zhang Lei. Deep convolutional and gated recurrent neural networks for sensor-based activity recognition [J]. *Journal of Electronic Measurement and Instrumentation*, 2020, 34(1): 1-9. (in Chinese)
- [10] Wang Y, Jiang X L, Cao R Y, et al. Robust indoor human activity recognition using wireless signals [J]. *Sensors*, 2015, 15(7): 17195-208.
- [11] Liu Xiwen, Chen Haiming. Wi-ACR: a human action counting and recognition method based on CSI [J]. *Journal of Beijing University of Posts and Telecommunications*, 2020, 43(5): 105-111. (in Chinese)
- [12] De P, Chatterjee A, Rakshit A. PIR sensor-based AAL tool for human movement detection: modified MCP-based dictionary learning approach [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(10): 7377-7385.
- [13] Pourpanah F, Zhang B, Ma R, et al. Non-intrusive human motion recognition using distributed sparse sensors and the genetic algorithm based neural network[C]//2018 IEEE Sensors. IEEE, 2018: 1-4.
- [14] Sun Q, Hu F. Dual-mode binary thermal sensing for indoor human scenario recognition with pyroelectric infrared sensors[C]//2019 IEEE Sensors. IEEE, 2019: 1-4.
- [15] Guan Q, Li C, Qin L, et al. Daily activity recognition using pyroelectric infrared sensors and reference structures [J]. *IEEE Sensors Journal*, 2018, 19(5): 1645-1652.
- [16] Yang Y, Yang H L, Liu Z X, et al. Fall detection system based on infrared array sensor and multi-dimensional feature fusion [J]. *Measurement*, 2022, 192: 110870.
- [17] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [18] Cho K, Van M B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[DB/OL]. (2014-06-03). <https://arxiv.org/abs/1406.1078>.
- [19] Chung J, Gulcehre C, Cho K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [DB/OL]. <https://arxiv.org/abs/1412.3555>.
- [20] Srivastava N, Hinton G, Krizhevsk A, et al. Dropout: A simple way to prevent neural networks from overfitting [J]. *Journal of Machine Learning Research*, 2014, 15(1): 1929-1958.