

运用开端神经网络进行人体姿态识别

柴铎^{1,2}, 徐诚^{1,2}, 何杰^{1,2}, 张少阳^{1,2}, 段世红^{1,2}, 齐悦^{1,2}

(1. 北京科技大学计算机与通信工程学院, 北京 100083; 2. 材料领域知识工程北京市重点实验室, 北京 100083)

摘 要: 通过迁移深度学习神经网络在图像识别方面的经验, 提出了一种基于 Inception 神经网络和循环神经网络结合的深度学习模型 (InnoHAR), 该模型端对端地输入多通道传感器的波形数据, 利用 1×1 卷积对多通道数据的有机组合, 不同尺度的卷积提取不同尺度的波形特征, 最大池化过滤微小扰动造成的假阳性, 结合时间特征提取 (GRU) 为时序特征建模, 充分利用数据特征完成分类任务。相比已知最优的神经网络模型, 在识别准确度上有近 3% 的提升, 达到了 state-of-the-art 的水平, 同时可以保证低功耗嵌入式平台的实时预测, 且在网络结构组成上更加丰富, 具有更大的潜力和挖掘空间。

关键词: 人体姿态识别; 行为识别; 开端神经网络; 可穿戴传感器网络; 深度学习

中图分类号: TP391

文献标识码: A

Inception neural network for human activity recognition using wearable sensor

CHAI Duo^{1,2}, XU Cheng^{1,2}, HE Jie^{1,2}, ZHANG Shao-yang^{1,2}, DUAN Shi-hong^{1,2}, QI Yue^{1,2}

(1. School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China;

2. Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China)

Abstract: The experience from computer vision was learned, an innovative neural network model called InnoHAR (inception neural network for human activity recognition) based on the inception neural network and recurrent neural network was put forward, which started from an end-to-end multi-channel sensor waveform data, followed by the 1×1 convolution for better combination of the multi-channel data, and the various scales of convolution to extract the waveform characteristics of different scales, the max-pooling layer to prevent the disturbance of tiny noise causing false positives, combined with the feature of GRU helped to time-sequential modeling, made full use of the characteristics of data classification task. Compared with the state-of-the-art neural network model, the InnoHAR model has a promotion of 3% in the recognition accuracy, which has reached the state-of-the-art on the dataset we used, at the same time it still can guarantee the real-time prediction of low-power embedded platform, also with more space for future exploration.

Key words: human activity recognition, activity recognition, inception neural network, wearable sensor network, deep learning

1 引言

人体姿态识别是普适计算领域的一项重要课题, 在医疗看护、体育竞技和智能家居^[1-3]等领域都有着广泛应用。通过对人体部位和周围环境的观测试图

解决的问题, 实现对复杂人体动作和行为的判别^[1-9]。最近几年, 随着深度学习技术的快速发展和日益流行, 其在人体姿态识别领域也得到了广泛的应用^[5,6]。与传统的统计机器学习方法不同, 深度学习方法通过自动对复杂高维非线性特征的提取和充分的学

收稿日期: 2017-09-19

通信作者: 齐悦, qiyue@ustb.edu.cn

基金项目: 国家重点研发计划基金资助项目 (No.2016YFC0901303); 国家自然科学基金资助项目 (No.61671056, No.61302065, No.61304257, No.61402033); 北京市自然科学基金资助项目 (No.4152036); 天津市重大科技专项基金资助项目 (No.16ZXCXSF00150)

Foundation Items: The National Key R&D Program of China (No.2016YFC0901303), The National Natural Science Foundation of China (No. 61671056, No. 61302065, No. 61304257, No.61402033), Beijing Natural Science Foundation Project (No.4152036), Tianjin Special Program for Science and Technology (No. 16ZXCXSF00150)

习，使其在面对大量不同传感器源的复杂数据进行特征提取和分类时显得更加游刃有余。

深度学习给计算视觉带来的飞速发展使很多研究者寄希望于通过计算视觉解决人体姿态识别的问题。但是，基于计算视觉的人体姿态识别方案仍然有很多不足，例如：1) 复杂多变的背景对动作识别的干扰；2) 多个活动主体同时出现在画面中造成的定位、追踪和识别的困难；3) 人体姿态识别对于图像采集环境光线、亮度、对比度的苛刻要求；被关注对象被遮挡等都影响着基于计算视觉的人体姿态识别的发展和在实际场景中的应用。

可穿戴传感器以其易部署、高精度、低功耗等优点，被广泛应用于人体姿态识别和动作捕捉等应用中^[8,9]。少数研究者也对深度学习在基于可穿戴传感器的人体姿态识别中的应用进行了探索，但多数已有研究仅仅是把深度学习作为一个黑盒子，在数据上浅尝辄止。例如，文献[6,9]等仅使用浅层的卷积神经网络实现滑动窗口的特征提取。文献[9]在此基础上使用了 LSTM，加入了对人体姿态识别的时序性的考虑。上述工作借助 CNN 和 LSTM 对人体活动的波形数据进行特征提取，在已有的公开数据集上相比经典的特征工程+机器学习的方法有了大幅的提升，展现了深度学习的强大威力。但是，对于人体日常活动中不同持续时长跨度很大(a survey on human activity recognition using wearable sensors)，不同活动间没有明确分界，人体无意识活动等带来的误识别等可穿戴传感器的人体姿态识别领域中的相关问题少有涉及。

针对上述问题，本文构造了一个基于 Inception 神经网络和 GRU 结合的多层次神经网络结构模型，主要贡献如下。

1) 提出了一种全新的深度神经网络模型 InnoHAR。

2) 通过使用 1×1、1×3、1×5 等不同尺度大小的卷积核，对波形数据进行特征提取和拼接，实现对不同持续时长的人体姿态特征提取。

3) 通过使用池化层对人体无意识抖动带入的干扰噪声的过滤，以降低姿态识别的误识别率。

4) 通过不同尺度的卷积层串联卷积和池化层的串联以及多次的非线性激活，使上述优点被叠加放大，使单个人体姿态的高维特征更易被提取，大大降低了人体姿态间没有明显分割导致的干扰和误识别。

5) 通过对比实验，证明了盲目的神经网络层数

的叠加和参数的堆积，除了成倍地提高训练和验证的时间，降低了学习和预测的效率，并不能提升识别的准确率，相反可能因为网络过深而数据量不足，导致网络结构难以被充分训练，进而导致预测结果的下降。

2 方法与设计

深度神经网络是深度学习方法的基础，其中卷积神经网络和循环神经网络是深度神经网络主流的两类网络结构。

2.1 卷积神经网络

卷积神经网络(CNN)的每一个神经元与前一层的局部接受域相连——它的功能就像是一个滤波器，然后通过非线性函数激活，如式(1)所示。

$$a_{i,j} = f\left(\sum_{m=1}^H \sum_{n=1}^K W_{m,n} x_{i+m,j+n} + b\right) \quad (1)$$

其中， $a_{i,j}$ 是相应的激活， f 是非线性函数， W 是 $H \times K$ 的卷积核的权重矩阵， b 是偏置值， x 表示连接到神经元 (i,j) 的上层神经元的激活。带有几个卷积层的 CNN，能够学习对数据的分层表示，更深层的卷积层用更抽象的方式表征数据。

2.2 长短期记忆

长短期记忆(LSTM)的输入被送入不同的门，包括输入门、输出门和遗忘门，分别代表对信息长期、远期和近期的记忆和控制。每一个 LSTM 单元激活通过式(2)计算。

$$a_t = \sigma(w_{ih}x_t + w_{hh}a_{t-1} + b) \quad (2)$$

其中， a_t 表示 t 时刻的激活， a_{t-1} 表示 $t-1$ 时刻的激活， σ 是非线性函数， w_{ih} 是输入层与隐藏层连接矩阵， w_{hh} 是隐层节点连接到自身的连接矩阵， b 是偏置值。

LSTM 的细胞状态每经过一个时间步长 t 都会被更新一次，计算式如下

$$i_t = \sigma_i(w_{xi}x_t + w_{hi}a_{t-1} + b_i) \quad (3)$$

$$f_t = \sigma_f(w_{xf}x_t + w_{hf}a_{t-1} + b_f) \quad (4)$$

$$o_t = \sigma_o(w_{xo}x_t + w_{ho}a_{t-1} + b_o) \quad (5)$$

$$c_t = f_t c_{t-1} + i_t \sigma_c(w_{xc}x_t + w_{hc}a_{t-1} + b_c) \quad (6)$$

$$a_t = o_t \sigma_a(c_t) \quad (7)$$

其中， i 、 f 、 o 分别代表输入门、遗忘门和输出门。 a 表示细胞的激活。 σ 表示非线性激活函数。 x 表示存储细胞的输入。 w 表示单元间的连接矩阵。 b

是偏置值。

在语音识别和机器翻译等领域,以 LSTM 为代表的 RNN 有着广泛的应用^[5,8~10]。

2.3 GoogLeNet 的 Inception module 模块

基于 Inception 模块的 GoogLeNet 是 Google 于 2014 年下半年提出的创新性的网络结构。该网络结构不局限于传统的序贯模型——层与层之间线性的连接起来,而是首次提出 Inception 模块。如图 1 所示,GoogLeNet 的 Inception 模块在得到上一层的结果作为输入后,从左至右依次进入 1 个 1×1 的卷积、1 个 1×1 的卷积串联 3×3 的卷积、1 个 1×1 的卷积串联 5×5 的卷积和 1 个 3×3 的最大池化层串联 1×1 的卷积。随后将上述子模块计算得到的结果进行拼接,输入下一层。

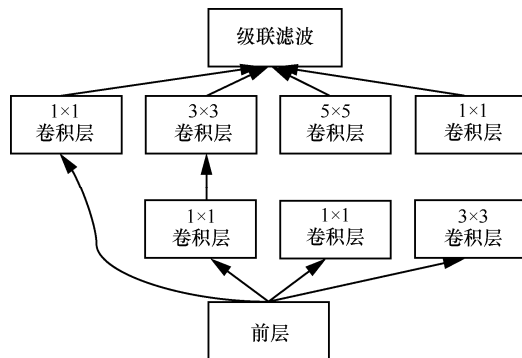


图 1 GoogLeNet 的 Inception 模块

在此之前,最直接提高深度神经网络性能的方法是增加其规模,包括通过增加层数以增大深度,通过增加每一层的节点数以增加宽度。这是训练高质量模型最简单安全的方法⁰,特别是对于给定的大规模标签数据集。然而这种简单的解决方法有两大缺陷。一是更大的网络规模往往意味着更多的参数,这使扩大后的网络更易过拟合,特别是当训练集中的标签样例有限的时候。另一个缺陷则是一味地增加网络大小的缺陷使计算资源需求的指数级的暴增。

本文采用了 Inception 模块,带来了多种被验证有效的网络构建技巧。例如, 1×1 的卷积网络实现数据的降维和多通道数据的线性组合和池化层的广泛应用实现降维、关键特征提取和滤波等作用。

3 系统结构

InnoHAR 深度神经网络模型的网络结构如图 2 所示。在数据输入后,首先通过 4 个类 Inception 模

块,其中依据本文研究的 1 维时序数据特征,使用 1×1 、 1×3 、 1×5 的卷积和池化层,具体的类 Inception 网络结构在第 3.1 节介绍。同时在 3、4 两个类 Inception 模块后,本文同样接最大池化层,帮助网络更好地消除噪声扰动带来的误判。最后,输出结果通过 2 个 GRU 层,使模型能更好地提取序列的时序的前后依赖关系。

3.1 空间特征提取

在传感器波形数据的特征信息提取上,本文借鉴第 2.3 节中 GoogLeNet 的 Inception 模块的构造,实现了在 Opportunity 数据集上的类 Inception 模块的构建,如图 2 所示。在每个类 Inception 模块中,本文同样使用一个 1×1 的卷积核直接将多通道信息的组合非线性激活后传入下一层,2 个分别为 1×3 和 1×5 的卷积核各自级联在一个 1×1 的卷积核之后,为整个模型提取不同尺度的特征信息,同时与只经过 1×1 卷积的输出拼接,还产生了类似 ResNet (deep residual learning for image recognition) 残差连接的效果。同时,还有一个 1×3 的池化层后级联一个 1×1 的卷积核来提供特征增强和滤波后的特征信息。这些类 Inception 模块中的子模块会分别使用 ELU 作为非线性的激活函数。随后这些子结构会被拼接起来,并再次经过 ELU 的非线性激活,输出到下一层。

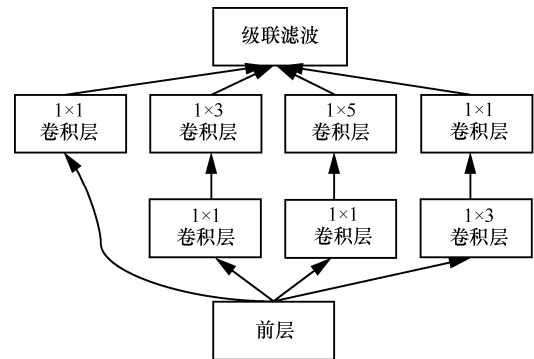


图 2 在 Opportunity 数据集上的类 Inception 模块

3.2 时间特征提取: GRU

在时序特征的提取上,参考文献[9]的工作,同样选择两层循环层进行时序特征的提取,为了实验的对比分析。经验表明, vanilla RNN 因为会存在梯度消失,即后面时间的节点对于前面时间的节点感知力下降的问题^[10],而在多个实验和作者的实验中,GRU 和 LSTM 在处理长序列的依赖问题上都有相似的相较 vanilla RNN 更好的表现,而 GRU 在

时间效率上的表现则更加出色, 基于第 3.1 节中相对复杂的网络结构, 在这里选用 GRU 作为循环层的具体实现, 使整个网络结构在预测性能和效率上同时带来令人满意的效果。图 3 中黑色的网络结构为池化层, 而灰色的网络结构即为 GRU 层。

3.3 模型实现

本文使用 Keras 2^[11]来构建网络结构。Keras 是一个高层次的神经网络 API, 由 Python 编写并可选 Tensorflow 或 Theano 作为后端。本文在实验中选择 Tensorflow 作为后端, 在 GPU 上运行。

4 实验设计与结果分析

4.1 Opportunity 数据集

Opportunity activity recognition dataset^[12]是一个在丰富传感器的环境中收集得到的人体自然活动的公开数据集。该数据集是对 4 个志愿者早晨日常活动的记录, 其中包含不同种类放置在包括人体不同部位和物体上的不同传感器信息。

4.2 预处理

为了尽可能减少前期的预处理工作, 实现端到端的人体姿态识别模型, 同时保持与已有工作的数据一致, 对于 Opportunity Challenge 的数据集, 本文同样采用全部的人体各部位的传感器数据, 共 113 个通道的数据, 并同文献[9]使用相同的固定长度 24 行为一个滑动窗口, 每次滑动 12 行, 进行数据分割, 保证全部数据的都被完整重叠。共分割得到 9 894 条数据用作测试。本文通过线性拟合填充传感器的缺失值, 每个传感器通道归一化到[0,1]。这是一个包含 null 类, 共 18 类的分类问题。

4.3 性能测试

自然场景下测得的人体活动数据集类与类之间是不平衡的。某些类可能包含大量的样本, 某些类只有少量的样本。Opportunity dataset 的手势是不平衡的, Null class 占的比重超过 70%。准确率

$$Precision = \frac{TP}{TP + FP}, \text{ 召回率 } Recall = \frac{TP}{TP + FN},$$

TP 、 FP 分别表示真正类和假正类, FN 表示假负类。为了解决类间的不平衡, 按每类的样本比例进行加权

$$F_1 = \sum_i 2w_i \frac{precision_i \cdot recall_i}{precision_i + recall_i}$$

其中, $w_i = \frac{n_i}{N}$ 为第 i 类样本所占的比例, n_i 表示第 i 类的数量, N 表示总样本的数量。

4.4 结果与分析

本文首先使用第 4.2 节预处理后产生的数据对模型的人体姿态识别准确程度进行评估。

DeepConvLSTM 是该数据集上之前的 state-of-the-art 模型, 本文在参照了文献[9]和 Github 托管的代码, 使用 Keras 2 重写, 尽可能完整地还原了文献[9]的训练的各个参数, 以减小不同语言框架等对实验结果的分析。本文对第 3 节中介绍的模型进行了评估, 同时为了展示池化层在降维、关键特征提取和滤波等方面的能力, 加入了对比实验, 将第 3 节中 Inception 模块内的池化层替换为 1×12 的卷积层以尽可能保证该模型与带有池化层的模型具有相同的特征量级。实验结果如图 4 所示, 由图 4 可知, 相比文献[9]的模型, 本文的基于类 Inception 模块的 InnoHAR 模型具有更加出色的表现, 在总体表现上, 继续有约 3% 的提升, 而即未包含池化层的模型, 相比序贯的 CNN^[7]与 LSTM 的叠加也有了 1.5% 的提升, 这表明了类似 Inception 这样的 Networkin Network 的不同组合所拥有的潜力和池化层在类似人体姿态识别这类噪音复杂的波形数据上拥有的不可忽视的去噪和这类复杂架构的抑制过拟合的能力。同时, 本文也测试了利用更简单、更安全的传统方法进行模型性能的提升。

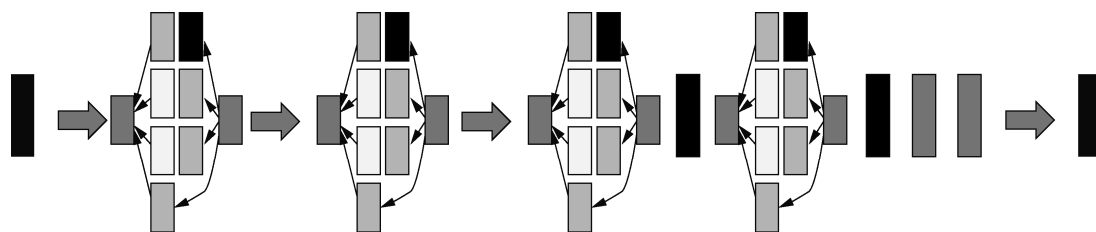


图3 面向人体姿态识别的开端神经网络系统结构

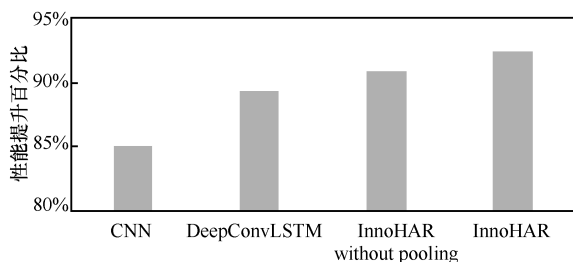


图 4 不同网络下实验结果

本文对于文献[9]的模型进行改造,实验结果如图 5 所示。在 DeepConvLSTM with 5 CNN layers 将模型原本的 4 层 CNN 增加 1 层与原卷积核大小相同的 CNN 层,变为 5 层,而在 DeepConvLSTM with bi-LSTM, 本文将模型中原来的单向 LSTM 变为每层分别有正反双向的 2 个 LSTM 结构,并将 2 个 LSTM 结构拼接输入下一层。在增加了一个 CNN 层试图简单提升模型的深度时,发现原本模型的识别准确度不仅没有提升,反而有一定幅度的下降,推测因为一个较大的 CNN 层包含着大量的参数,而 CNN 层的层与层之间采用着全连接的形式相连,这样每增加一层可能导致参数的指数上升。在数据量不够庞大的情况下,末层的梯度变化无法充分地传递给前面的层级,造成了整个模型的过拟合,反而影响了模型的准确程度。相比而言,在将原模型改造为双向 LSTM 后,模型的准确度有了小幅的提升,但是 LSTM 作为时序进行建模的网络层,后一个节点的计算必须等待前一个节点计算完成,方可开始计算,因而较难通过 GPU 并行化,因而在实际的训练过程中,常常占用一大部分的训练时间,而收益却相对没有那么明显。

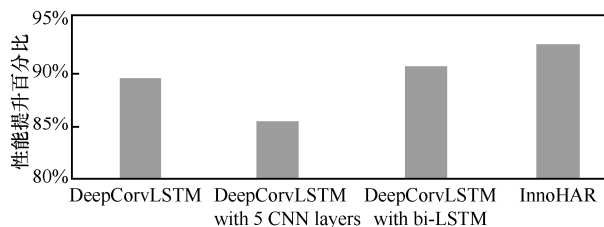


图 5 改进网络下的实验结果

4.5 案例研究

本文也对 InnoHAR 模型对测试数据的预测做了部分案例分析和一些统计(见表 1)。例如在前 9 000 个测试数据中,共有 649 个错误的分类结果,其中,原本为正常的类别动作标记,而被误判为 null 类的预测有 358 个,占到总的错误比例的 55.16%。对于这个问题,本文推测有两类原因,一个是训练

和测试数据本身的严重不平衡,导致模型本身有趋向于判为 null 分类的倾向,对于这个问题,本文认为可以在未来的工作中通过数据增量的办法或通过先训练二分类器进行 null 和非 null 的预测,而后再进行分类等办法尝试解决;另一个原因则是如 Opportunity Challenge 数据集所提到的,在数据以毫秒为单位标记过程中,难免会出现 overfill 和 underfill 的情况,这个趋向于数据集本身的特性,较难避免。

本文也注意到,在一些预测中,模型会偶然地在连续的 null 类别序列或其他类别的序列中给出单个或极少数个不同类别的预测,而在人类本身观测中可以轻易地推测出他们可能受到噪音信息的干扰产生了误判,这是因为人们清楚地认识到,人的动作姿态变化通常是一个连续非高频变化的过程,不会在几十毫秒内产生“状态抖动”,对此,本文建议在未来的工作中,对于 softmax 产生的各个类别的置信度,可以不必直接根据置信度最大的类别给出分类结果,而通过例如简单地依据当前预测项前后的预测结果加权和阈值处理,便可得到更好的效果。

4.6 耗时分析

在运行时间和效率的测试上,没有像以往的实验简单地统计运行时间而后给出结果。人体姿态识别模型的应用流程通常是先通过研究者携带传感器收集和标记数据训练模型,而后将训练的模型导入嵌入式系统中,嵌入式系统通过读取传感器当前的实时数据,实时给出姿态的预测。因而笔者认为,无论模型的精度高低,必须要考虑模型在嵌入式系统的实际运行情况。

本文在实验中采用的嵌入式系统是 MinnowBoard Turbot Dual Core Board, MinnowBoard Turbot Dual Core Board 是 Intel 于 2015 年发布的小型嵌入式平台,它搭载了 Intel Atom E3826 处理器,拥有 1.46 GHz 主频,2 GB DDR3L 1067MT/s DRAM,是一款兼具性能和价格优势的平台选择。

本文在其上搭载了 Ubuntu 14.04 操作系统、Keras 2.0.3 和 Tensorflow 1.0.1 进行测试,将模型文件和权重文件导入系统,在系统开始预测计时时,使用前文的 InnoHAR 模型,在对之前的全部测试数据的预测中,模型耗时 152.02 s,完成了全部 9 894 个滑动窗口的预测。预测速度达到了 65.09 个/秒。而原测试数据以 30 Hz 进行数据记录,本文的模型可以在该平台上较轻松地实现实时姿态的预测。

表 1 InnoHAR 模型预测结果的混淆矩阵

案例	null	Open Door 1	Open Door 2	Close Door 1	Close Door 2	Open Fridge	Close Fridge	Open Dishwasher	Close Dishwasher	Open Drawer 1	Close Drawer 1	Open Drawer 2	Close Drawer 2	Open Drawer 3	Close Drawer 3	Clean Table	Drink from Cup	Toggle Switch
null	7724	13	6	11	3	57	17	38	33	9	18	8	4	16	8	48	116	35
Open Door 1	4	40	0	8	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Open Door 2	2	0	88	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
Close Door 1	0	5	0	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Close Door 2	2	0	1	0	76	0	0	0	0	0	0	0	0	0	0	0	0	0
Open Fridge	14	0	0	0	0	143	8	1	0	0	0	0	0	0	0	0	0	0
Close Fridge	32	0	0	0	0	26	135	0	0	0	0	0	0	0	0	0	0	0
Open Dishwasher	9	0	0	0	0	1	0	60	3	0	0	0	0	0	0	0	3	0
Close Dishwasher	17	0	0	0	0	1	0	0	40	0	0	0	1	0	0	0	0	0
Open Drawer 1	2	0	0	0	0	0	0	1	0	24	6	10	3	1	0	0	0	0
Close Drawer 1	5	0	0	0	0	0	0	0	1	4	18	1	2	0	0	0	0	0
Open Drawer 2	3	0	0	0	0	0	0	0	0	0	0	18	1	4	0	0	0	0
Close Drawer 2	3	0	0	0	0	0	0	0	0	1	0	0	15	1	6	0	0	0
Open Drawer 3	6	0	0	0	0	0	0	0	0	0	0	1	0	42	13	0	0	0
Close Drawer 3	14	0	0	0	0	0	0	0	0	0	0	2	0	3	34	0	0	0
Clean Table	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	51	0	0
Drink from Cup	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	197	0
Toggle Switch	10	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	70

5 结束语

借鉴图像识别中的 Inception 网络结构，通过将不同尺度的卷积核进行级联和与池化层的拼接，概念性地提出了 InnoHAR 基于传感器网络的人体姿态识别的创新模型。该模型相比传统的机器学习模型平均有近 15%的提升，即使相比之前构建的识别准确度最高的 DeepConvLSTM 也有了 3%的提升，达到了在该数据集上的 state-of-the-art。实验中也证明了本文的创新结构相比传统的序贯叠加深度学习模型鲁莽地层数增加的结构具有更大的潜力。

对于未来的工作方向，首先会继续调整网络结构，包括大小核的尺寸、连接方式等；也会针对文中提到的数据不平衡的问题和最终输出的类别的选择做进一步的探索；同时也会考虑借助迁移学习将该模型应用到本地的数据中，充分利用公开测试集上积累的丰富的数据资源和训练结果，使模型发挥更大的作用。

参考文献：

[1] XU C, HE J, ZHANG X T, et al. Detection of freezing of gait using

template-matching-based approaches[J]. Journal of Sensors, 2017, 2017(2):1-8.

[2] XU C, HE J, ZHANG X T, et al. Geometrical kinematic modeling on human motion using method of multi-sensor fusion[J]. Information Fusion, 2017.

[3] MARGARITO J, HELAOUI R, BIANCHI A M, et al. User independent recognition of sports activities from a single wrist-worn accelerometer: a template-matching-based approach[J]. IEEE Transactions on Bio-medical Engineering, 2015, 63(4):788-796.

[4] ROY P C, GIROUX S, BOUCHARD B, et al. A possibilistic approach for activity recognition in smart homes for cognitive assistance to alzheimer’s patients[M]// Activity Recognition in Pervasive Intelligent Environments. Atlantis Press, 2011:33-58.

[5] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// Computer Vision and Pattern Recognition. 2014:1-9.

[6] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Computer Vision and Pattern Recognition. 2015: 770-778.

[7] BENGIO Y, LECUN Y. Convolutional networks for images, speech, and time-series[J]. Convolutional Networks for Images, Speech, and Time Series, 1995.

[8] GERS F A, SCHRAUDOLPH N N, SCHMIDHUBER, J, et al. Learning precise timing with lstm recurrent networks[J]. Journal of Machine Learning Research, 2003, 3(1):115-143.

[9] ORDÓÑEZ F J, ROGGEN D. Deep convolutional and LSTM Recurrent neural networks for multimodal wearable activity recognition[J]. Sensors, 2016, 16(1):115.

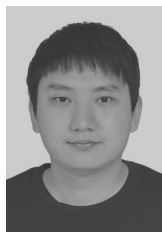
[10] GERS F A, SCHRAUDOLPH N N, SCHMIDHUBER, J, et al. Learn-

ing precise timing with lstm recurrent networks[J]. Journal of Machine Learning Research, 2003, 3(1):115-143.

- [11] XU C, HE J, ZHANG X T, et al. Toward Near-Ground Localization: Modeling and Applications for TOA Ranging Error[J]. IEEE Transactions on Antennas & Propagation, 2017, 65(10):5658-5662.

- [12] CHAVARRIAGA R, SAGHA H, CALATRONI A, et al. The opportunity challenge: a benchmark database for on-body sensor-based activity recognition[J]. Pattern Recognition Letters, 2013, 34(15): 2033-2042.

作者简介:



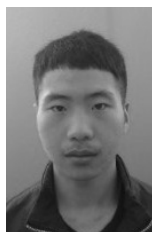
柴铎 (1995-), 男, 陕西府谷人, 北京科技大学工程师, 主要研究方向为聊天机器人、机器翻译等深度学习在自然语言处理领域的应用。



徐诚 (1988-), 男, 辽宁开原人, 北京科技大学博士生, 主要研究方向为室内定位、无线网络、模式识别等。



何杰 (1983-), 男, 浙江台州人, 博士, 北京科技大学副教授, 主要研究方向为室内定位、无线网络、模式识别等。



张少阳 (1991-), 男, 河北石家庄人, 北京科技大学硕士生, 主要研究方向为模式识别等。

段世红 (1973-), 女, 山西太原人, 博士, 北京科技大学副教授, 主要研究方向为计算机软件、无线网络、模式识别等。

齐悦 (1975-), 女, 辽宁沈阳人, 博士, 北京科技大学副教授, 主要研究方向为计算机软件、无线网络、模式识别等。