

基于残差时域注意力神经网络的交通模式识别算法

刘世泽¹, 朱奕达², 陈润泽², 罗海勇^{3*}, 赵方², 孙艺², 王宝会¹

(1. 北京航空航天大学 软件学院, 北京 100191; 2. 北京邮电大学 计算机学院(国家示范性软件学院), 北京 100876;

3. 中国科学院 计算技术研究所, 北京 100190)

(* 通信作者电子邮箱 yhluo@ict.ac.cn)

摘要: 交通模式识别是用户行为识别中的一个重要分支,其目的是对用户所处的交通模式进行准确判断。针对现代智慧城市交通系统对在移动设备环境下精准感知用户交通模式的需求,提出了一种基于残差时域注意力神经网络的交通模式识别算法。首先,通过具有较强局部特征提取能力的残差网络提取传感器时序中的局部特征;然后,采用基于通道的注意力机制对不同传感器特征进行重校准,并针对不同传感器的数据异构性进行注意力重校准;最后,利用具有更广感受野的时域卷积网络(TCN)提取传感器时序中的全局特征。采用数据丰富度较高的宏达通讯(HTC)交通模式识别数据集来对已有的交通模式识别算法和所提出的残差时域注意力模型进行评估,实验结果表明,所提出的残差时域注意力模型在对现代移动嵌入式设备的计算开销友好的前提下具有高达96.07%的准确率,且对单一类别均具有高于90%的召回率与精确率,验证了该模型的准确性与鲁棒性。所提模型可以作为一种支持移动智能终端运算的交通模式识别应用于智能交通出行、智慧城市等领域。

关键词: 时域卷积网络;交通模式识别;残差网络;注意力机制;深度学习

中图分类号: TP183 **文献标志码:** A

Traffic mode recognition algorithm based on residual temporal attention neural network

LIU Shize¹, ZHU Yida², CHEN Runze², LUO Haiyong^{3*}, ZHAO Fang², SUN Yi², WANG Baohui¹

(1. College of Software, Beihang University, Beijing 100191, China;

2. School of Computer Science (National Pilot Software Engineering School),
Beijing University of Posts and Telecommunications, Beijing 100876, China;

3. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Traffic mode recognition is an important branch of user behavior recognition, the purpose of which is to identify the user's current traffic mode. Aiming at the demand of the modern intelligent urban transportation system to accurately perceive the user's traffic mode in the mobile device environment, a traffic mode recognition algorithm based on the residual temporal attention neural network was proposed. Firstly, the local features in the sensor time sequence were extracted through the residual network with strong local feature extraction ability. Then, the channel-based attention mechanism was used to recalibrate the different sensor features, and the attention recalibration was performed by focusing on the data heterogeneity of different sensors. Finally, the Temporal Convolutional Network (TCN) with a wider receptive field was used to extract the global features in the sensor time sequence. The data-rich High Technology Computer (HTC) traffic mode recognition dataset was used to evaluate the existing traffic mode recognition algorithms and the residual temporal attention model. Experimental results show that the proposed residual temporal attention model has the accuracy as high as 96.07% with friendly computational overhead for mobile devices, and has the precision and recall for any single class reached or exceeded 90%, which verify the accuracy and robustness of the proposed model. The proposed model can be applied to intelligent transportation, smart city and other domains as a kind of traffic mode detection for supporting mobile intelligent terminal operation.

Key words: Temporal Convolutional Network (TCN); traffic mode recognition; Residual Network (ResNet); attention mechanism; deep learning

收稿日期: 2020-11-04; 修回日期: 2021-03-29; 录用日期: 2021-04-06。

基金项目: 国家重点研发计划项目(2019YFC1511400); 国家自然科学基金资助项目(61872046); 北京市自然科学基金资助项目(4212024); 北京邮电大学提升科技创新能力行动计划项目(2019XD-A06); 北京市自然科学基金-海淀原始创新联合基金资助项目(L192004); 河北省重点研发计划项目(19210404D); 内蒙古自治区关键技术攻关计划项目(2019GG328); 移动计算与新型终端北京市重点实验室开放课题。

作者简介: 刘世泽(1988—), 男, 辽宁抚顺人, 硕士, 主要研究方向: 大数据挖掘、智能感知; 朱奕达(1996—), 男, 浙江慈溪人, 博士研究生, 主要研究方向: 城市计算、交通模式识别; 陈润泽(1996—), 男, 甘肃白银人, 博士研究生, 主要研究方向: 城市计算、交通模式识别、位置语义挖掘; 罗海勇(1967—), 男, 湖北麻城人, 副研究员, 博士, CCF 会员, 主要研究方向: 移动智能、普适计算; 赵方(1968—), 女, 河南开封人, 教授, 博士, CCF 会员, 主要研究方向: 移动互联网、大数据挖掘; 孙艺(1979—), 男, 山东菏泽人, 高级工程师, 硕士, 主要研究方向: 移动互联网、大数据挖掘; 王宝会(1973—), 男, 江苏滨海人, 教授级高级工程师, 硕士, 主要研究方向: 软件工程。

0 引言

随着移动智能终端设备的飞速发展,智能手机已经逐渐成为人们日常生活中的必需品。搭载丰富内置传感器的智能终端设备具备多维度感知用户所处上下文环境信息的能力。近年来,通过捕获用户的上下文环境信息来感知用户的行为模式^[1-3]成为了普适计算中的研究热点。用户行为识别通常基于两种数据源来完成,包括基于视觉信息的用户行为识别和基于轻量级传感器的用户行为识别。基于视觉的方法需要外部系统捕捉图像特征,对环境的要求较高并需要良好的光线条件。基于轻量级传感器的方法不依赖基础设施建设,具有较强的普适性。

交通模式识别是用户行为识别中一个重要的检测分支,其目的是识别用户当前的交通模式,即对用户所处的交通模式进行准确判断。用户在出行过程中常常使用不同的交通工具,在使用不同交通工具的过程中,常常会具有不同的出行需求。由于这些需求需要移动智能终端预先获得用户所处的交通工具状态,即交通模式,交通模式识别问题是一个在多个领域起到至关重要作用的基础问题,不同场景下基于精准的交通模式感知,才能够提供围绕用户出行的一系列服务,如交通状况预测^[4-5]、用户轨迹挖掘^[3,6-7]、智能出行推荐^[8]等。交通模式识别需要用户随身携带数据采集设备,基于视觉的方法进行交通模式检测难以在实际应用场景进行部署,所以大部分的研究在围绕交通模式识别问题上使用基于传感器数据的方法^[1,9-14]。交通模式识别在各行各业都有着较为丰富的应用需求:在外卖闪送等实际业务场景下,使用交通模式识别技术能够帮助大数据平台进一步精确识别订单的配送情况,平台对配送人员精确感知能够避免订单的异常配送状态,例如配送人员提前完成订单,平台难以对延迟配送进行判定,配送人员在配送过程中出现特殊状况,如交通事故、人身伤害等状况,更加精确的交通模式感知能够帮助保护平台用户的利益与配送人员的安全;在移动设备人机交互或移动地图智能规划的业务场景下,移动设备精准感知用户的交通模式能够帮助智能终端对用户的出行模式进行智能规划,向用户提供更加人性化的出行建议与出行规划;在用户轨迹与交通状况进行大数据分析的智慧城市业务场景下,大量用户的实时交通模式数据可以进一步帮助城市进行交通状况分析,及时规划公共交通工具的智能调度,提高城市公共交通资源调度的科学性,避免交通资源的浪费以及紧缺。

考虑到不同交通模式下传感器信号变化特点不同,文献[2,15-18]等在基于传感器数据的交通模式识别方法中人工提取多源传感器信号特征。他们将提取的特征输入到传统机器学习算法中完成交通模式的预测,例如支持向量机^[11,6]、决策树^[9,19-20]、集成学习^[16-18,20]等。然而这些基于传统机器学习算法的方法主要取决于手工提取特征的有效性,这既耗时又需要领域相关知识。

近年来,深度学习的研究引起了广泛关注,由于深度学习不依赖特有的领域知识并具有从大规模原始数据中自动提取特征的能力。很多研究者^[2,11-13]采用深度学习的方法来完成交通模式的预测。卷积神经网络可以通过叠加几个卷积运算自动创建抽象的层次结构特征。文献[1,9,12]利用卷积神经网络从加速度、陀螺仪、地磁等轻量级传感器分别提取特征后

进行分类。传感器信号随时间序列变化,时间序列神经网络能够合理地模拟时间序列问题中的时间依赖性。文献[14,21]提出利用循环卷积网络和长短期记忆递归神经网络提取传感器的时域特征进行交通模式识别。

但是现有交通模式识别方法中依旧存在不足:卷积神经网络存在只能提取局部特征的问题,时间序列建模的循环神经网络由于其模型结构复杂并且不能并行计算,存在模型参数量大、训练时间长的的问题;此外,没有对分类特征进行特征重要性校准。为了解决以上问题,本文提出了残差时域注意力神经网络进行交通模式识别,能够达到准确性与鲁棒性的平衡,满足了对现代移动嵌入式设备友好的计算开销下的实际交通模式识别性能需求。

1 相关工作

近十年来围绕机器学习算法的创新与应用得到了巨大发展,并在普适计算领域得到充分应用。在交通模式识别领域,学者们利用机器学习算法提供了许多解决方案。近十年来随着服务器计算能力的增强,使得深度学习算法快速发展,深度学习方法受到更多研究者的关注并为交通模式识别提供了新的解决方案。

1.1 基于移动传感器数据的交通模式识别

在交通模式识别问题的研究中,研究者往往关注交通模式识别的准确性与实时性,交通模式识别的过程需要在模型的准确度与普适性之间达到一定的平衡。在移动传感设备普及之前往往难以达到这样的效果,然而随着移动智能终端的逐渐普及,采集交通模式识别所需数据的硬件平台已趋于成熟。交通模式识别问题目前存在基于不同数据源建立的解决方案,包括基于全球导航卫星系统(Global Navigation Satellite System, GNSS)的解决方案与基于低功耗传感器的解决方案。近几年来随着低功耗传感器在智能手机中的大量普及,采用低功耗传感器数据的交通模式识别解决方案逐渐增多^[1,2,15,22-24]。

1.2 基于机器学习算法的交通模式识别

在深度学习算法得到普遍应用之前,许多研究者提出多种基于传统机器学习算法的交通模式识别解决方案,但由于机器学习算法需要研究人员手动提取特征,这些特征依赖于领域知识,所以在识别精度上存在较大提升空间。文献[6]中提出了一种从GNSS轨迹数据提取轨迹特征后,将特征送入到支持向量机分类器进行交通模式识别的方法,从而提供基于轨迹的交通模式分析。文献[20]利用GNSS数据提取卫星相关特征并基于贝叶斯网络、决策树、随机森林、朴素贝叶斯以及多层感知机进行交通模式识别。

由于使用GNSS传感器采集数据会导致移动终端功耗变大,很多研究者开始研究基于低功耗传感器数据进行交通模式识别。文献[1]提出了一种软硬件结合系统,通过搭建包含加速度传感器、地磁传感器以及陀螺仪的低功耗设备平台采集数据后,从而提取少量统计学特征,利用支持向量机(Support Vector Machine, SVM)作为分类器进行交通模式识别。文献[9]中利用智能手机中的加速度传感器、陀螺仪以及磁力传感器作为数据源,利用Filter相关性度量(Correlation-based Feature Selection, CFS)算法对传感器数据时域、频域特征进行提取并采用分层的轻量级梯度提升器(Light Gradient

Boosting Machine, LightGBM)算法对交通模式进行识别。文献[17]通过集成分类器XGBoost(Extreme Gradient Boosting, 极端梯度提升)与多层感知机的方式,基于加速度传感器、陀螺仪、地磁传感器以及气压传感器数据提取统计学、频域等特征进行交通模式识别。此外,还有研究人员利用高功耗GNSS数据和低功耗传感器一起进行交通模式的感知。文献[19]提出了采用Wi-Fi、蓝牙、GNSS、加速度、重力加速度以及陀螺仪等传感器提取用户轨迹特征和时域特征并使用随机森林算法进行交通模式识别。

1.3 基于深度学习算法的交通模式识别

随着深度学习算法的快速发展,深度学习模型具备不需要领域知识即可自动提取特征的能力。在交通模式识别问题上,深度神经网络也逐渐显现出其重要价值。研究者逐渐受到深度学习模型的启发,在交通模式识别问题上提出很多新解决方案。

卷积神经网络在传感器信号特征提取过程中起到了一定作用,不少研究者基于卷积神经网络进行了交通模式识别问题上的探索。文献[12]提出了拥抱网络(EmbraceNet)来解决交通模式识别问题,对每个传感器分别利用卷积神经网络进行特征提取,在提取出的复杂时域与频域信号特征中利用一种基于多项式分布权重构建的融合层对特征进行融合后完成分类。文献[11]提出采用稠密网络(DenseNet)^[25]进行多个传感器时域特征的提取,使用线性层进行特征融合来完成交通模式识别。

循环神经网络在时域信号特征提取上具有鲜明的作用,基于传感器数据的交通模式识别往往利用的是时间序列数据,许多研究者利用循环神经网络以及其变体进行时域特征的进一步挖掘。文献[10]提出将传感器数据送入循环神经网络进行时域特征提取,并利用提取出的时域特征进行分类预测。一些研究者^[14,21]使用长短期记忆模型进行基于历史信息特征提取,长短期记忆模型是循环神经网络的变体,引入了长期记忆与短期记忆机制,能够更好地对时域信息中的长周期特征进行挖掘。此外,文献[13]利用对抗自编码器针对时域传感器数据进行了交通模式识别。

然而上述工作一方面缺乏对时域特征的深层次挖掘,另一方面循环神经网络在计算复杂度与内存占用上仍然存在较大提升空间,许多工作仅关注于提高模型的准确率而使用大量的计算资源进行交通模式识别,难以在移动智能终端进行部署与优化,随着移动设备计算能力的不断提高,移动设备已经达到部署深度学习或机器学习算法的能力,但高计算开销也极大影响着实际模型端到端部署时实际运行的性能。此外,上述工作缺乏模型对细粒度特征的重校准能力,导致不同传感器信号的特征空间难以进行融合并且难以体现具有更明显交通模式特征的传感器在实际识别过程中起到的重要性,也导致了部分传感器特征的浪费。所以本文提出残差时域注意力网络以解决上述问题,本文模型能够结合时域卷积网络(Temporal Convolutional Network, TCN)的低计算开销同时也更为关注重要的传感器特征,避免了较低计算开销要求下模型性能的损失,具有较好的鲁棒性与稳定性。

2 残差时域注意力交通模式识别模型

本文提出了一种基于残差时域注意力神经网络的交通模

式识别模型,如图1所示,多源传感器时间序列数据输入到该模型后,该模型先利用残差卷积神经网络对不同传感器进行局部感受野的特征提取。该模型在提取各传感器局部特征后将其输入到时域卷积神经网络中进一步提取全局感受野的深层特征。此外,局部特征对于交通模式识别也有一定的必要性,为了保证模型能够充分利用局部特征,本文借助残差连接的方式传递浅层特征到模型的分层,使时域卷积网络能够提高对重要特征进行感知的能力,而时域卷积网络的膨胀卷积结构在提取局部细粒度特征时具有局限性也能得到弥补。由于不同传感器特征对于交通模式识别的重要性不同,本文使用了一种基于特征通道间关系的注意力机制对模型提取的交通模式识别分类特征进行特征重校准,使卷积网络能够充分感知真正具有较高重要性的复杂高维特征,从而降低了神经网络在识别过程中对无效高维特征进行大量计算导致的计算开销浪费问题,便于算法模型控制其算力规模,同时也提高了计算模型的鲁棒性,达到了计算效率与计算性能的平衡。

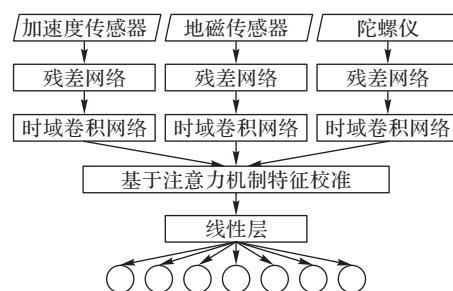


图1 本文模型结构示意图

Fig. 1 Schematic diagram of proposed model structure

2.1 基于注意力机制的残差卷积神经网络

深度神经网络是由多个非线性处理层堆叠而成的模型,卷积神经网络作为神经网络的一种,在提取小范围局部特征方面优势明显。但是随着网络层数的增加,梯度消失或梯度爆炸问题的出现导致深度神经网络在训练时难以收敛。残差卷积神经网络广泛应用于解决随着模型加深导致的梯度消失或梯度爆炸的问题,它使用的捷径连接将输入特征添加到输出特征上,从而提高浅层网络特征的流动性。残差连接能够确保模型中保持初始特征的恒等映射,在较深卷积网络之中依然能够保留原始特征,从而能够帮助模型在细粒度特征挖掘过程中获得更佳效果。针对时间序列传感器数据的交通模式识别,局部特征能够捕获交通模式在极小时间窗口内的变化规律,比如汽车突然的加速或者刹车、跑步中的摆臂等。本文提出的模型利用残差卷积神经网络对不同的传感器原始观测值进行局部特征的提取。

本文使用的残差神经网络由多个残差单元组合堆叠而成,本文使用的残差神经网络的框架如图2所示。给定输入数据,首先模型将输入的传感器数据经过卷积层、批处理标准化层和指数缩放线性单元(Scaled Exponential Linear Unit, SELU)处理之后,将处理得到的特征送入多个残差单元中进行处理,将最后一个残差单元中得到的局部特征经过批处理标准化层和池化层后送入到时域卷积网络模型中进行细粒度全局特征提取。

每个残差单元由卷积、批归一化和SELU堆叠而成。图3

展示了本文采用残差单元的示意图,其中第 l 个残差单元的输入为 x_l ,则其输出可形式化为执行式(1)中的数学计算。

$$x_{l+1} = f(W_s x_l + F(x_l, W_l)) \quad (1)$$

其中: $F(x_l, W_l)$ 为残差函数, W_l 为该残差函数的对应权重参数; $f(\cdot)$ 为非线性激活函数 SELU; W_s 为瓶颈层的权重参数,瓶颈层能够保证输入特征和输出特征的通道数一致。残差单元块中残差函数 $F(\cdot)$ 形式灵活可变,能够堆叠多个卷积层。

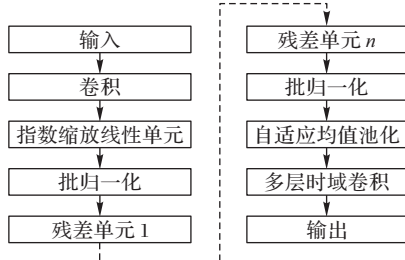


图 2 残差网络结构示意图

Fig. 2 Schematic diagram of ResNet structure

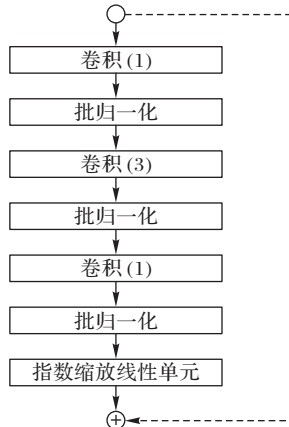


图 3 残差单元示意图

Fig. 3 Schematic diagram of residual unit

2.2 时域卷积网络

残差神经网络只针对单一传感器进行了局部的特征感知,单一传感器包括加速度、重力加速度以及地磁传感器提供的高维特征空间,只具有模型对单一环境上下文的局部感知,仅仅关注于用户的姿态、用户环境的磁场变化等单一变化,不同环境上下文之间的相互关联关系等隐藏特征仍需要进一步进行挖掘,因此本文引入了时域卷积网络来进行粗粒度的时序挖掘,同时也对不同传感器之间进行了粗粒度的融合感知。

循环神经网络具有较强的提取时序特征能力,其分支例如长短期记忆与门控循环神经网络进一步改进了循环神经网络的时序记忆范围,将较早的时序数据也能够充分利用起来。然而循环神经网络也具有一些显著的弱点,例如循环神经网络由于运算时依赖上下文输入的特征,需要进行密集计算导致模型引入大量的计算参数;其次,循环神经网络内存占用也较高,难以大规模并行处理。

针对以上问题,研究者提出了时域卷积网络^[26],时域卷积网络在较长的时间序列中占用较低的计算开销便可以充分挖掘全局特征,从而达到了理想的特征提取效果。本文提出的模型中使用时域卷积网络提取长时间窗口下交通模式的全局特征。时域卷积网络可以接收任意长度的时间序列数据,并

将其映射成任意长度的输出特征向量。本文模型中时域卷积网络的输入是由残差网络输出的传感器局部特征,然后利用时域卷积网络通过结构上的因果关系卷积充分挖掘时间序列上局部特征之间的潜在关系,从而保证在长时间序列下模型提取细粒度的全局特征。

本文提出的模型中的时域卷积网络由多个膨胀卷积块堆叠组成。膨胀卷积块是卷积神经网络的一个变种,虽然卷积神经网络能够对固定时间窗口的特征数据进行时序特征挖掘,但是由于传感器特征窗口本身即为连续的离散时间序列,通过卷积操作仅能够抽取在卷积核范围内的短期历史特征。为了能够在不改变卷积核大小的前提下扩大卷积核感受野,本文使用了膨胀卷积。对于一个特征窗口 $X \in \mathbf{R}^n$ 上第 s 个位置进行卷积操作 $f: \{0, 1, \dots, k-1\} \rightarrow \mathbf{R}$,则膨胀卷积为:

$$F(s) = (X *_{d} f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i} \quad (2)$$

其中, d 为膨胀系数,而 $x_{s-d \cdot i}$ 即为时间窗口中 s 位置之前的前序位置。在多层膨胀卷积结构中,第 i 层的膨胀系数设置为 $d_i = 2^{i-1}$ 。通过膨胀卷积,能够增大卷积核感受野范围,增强了模型对历史信息感知能力。通过增加膨胀卷积网络的深度,能够进一步加强膨胀卷积层对交通模式中历史信息更长时间尺度上的感知能力。在每次膨胀卷积之后本文引入批归一化和 SELU 进行特征非线性变化。对于传统时序模型某一时刻的计算需依赖于上一层的计算结束,但是在时域卷积模型中由于共享卷积核的参与,可以保证模型能够进行并行计算。此外,传统时间序列神经网络在使用时需要将每个时间步的信息都保存下来,从而占据大量内存,本文模型在每一层膨胀卷积中共享卷积核,内存占用更低。一个多层膨胀卷积块的操作过程如图 4 所示。

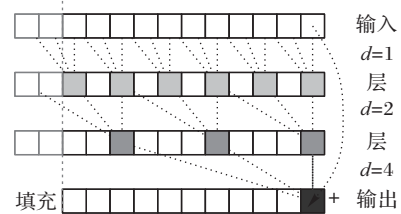


图 4 多层膨胀卷积结构示意图

Fig. 4 Schematic diagram of multi-layer dilated convolution structure

多层膨胀卷积块是本文提出的模型中时域卷积网络的基本结构,多层膨胀卷积的输出特征与输入特征具有相同的特征序列尺寸。为了避免训练过程中随着模型加深导致梯度消失的问题以及提高特征的流动性,本文模型在每个多层膨胀卷积块的输出层还加入了捷径连接,即将多层膨胀卷积块的输入特征与输出特征进行按位相加。在膨胀卷积块计算得到最终的输出特征之后,本文使用池化层对输出特征维度进行压缩并将压缩后的特征输入到下一个膨胀卷积块中。

2.3 特征重校准

在上述两种计算单元间进行特征传递的过程中,不同传感器特征从不同角度体现时间序列数据在交通模式上的变化规律,然而并非所有的特征都对交通模式分类任务起到相同作用。本文提出的模型在使用多维特征进行分类时考虑到不同特征之间的差异性,因此该模型引入收缩-激励注意力机

制(Squeeze-Excitation)^[27]对分类特征进行重校准,从而提高了交通模式的鲁棒性。特征重校准平衡了不同传感器特征空间中,由于单一传感器数据之间的差异性导致的特征空间差异性,在注意力机制的作用下,不同传感器通道之间得到了再次平衡,避免了机械地对异传感器间进行合并导致的统计差别针对不同特征通道的再平衡,充分发掘具有强重要性的特征通道能够在整个计算模型中发挥其代表性。

当不同传感器特征经由模型运算得到输出特征 W_{l+1} 后,本文利用式(3)~(5)进行基于通道间注意力机制的特征重校准。首先每个通道的特征经过全局压缩及拼接得到 d_c ,然后通过使用Sigmoid激活函数 σ 和SELU激活函数 δ 的门控机制对压缩得到的通道指示器 d_c 进行特征重要性的自动学习,最后将得到的通道特征收缩因子 s 对原始特征进行激励。

$$d_c = F_{\text{squeeze}}(x_{l+1}) = \frac{1}{H} I_c(i) \quad (3)$$

$$s = F_{\text{excitation}}(d, W) = \sigma(W_2 \delta(W_1 d)) \quad (4)$$

$$o_c = F_{\text{scale}}(i_c, s_c) = s_c \times i_c \quad (5)$$

在对不同传感器特征通道进行重校准后,对交通模式识别更有帮助的特征将被增强,不同交通模式下区分度较低的特征将被抑制,从而提高本文模型的识别准确性。

3 实验与结果分析

3.1 实验数据处理

本文采用数据丰富度较高的宏达通讯(data-rich High Technology Computer, HTC)提出的交通模式识别数据集^[18]进行实验,该数据集将交通模式划分为站立、行走、跑步、自行车、摩托车、轿车、公交以及轨道交通共8种类别。在实验过程中,本文随机抽取数据集中70%的部分作为训练集,30%的部分作为测试集。本文将使用6种具有代表性的模型与本文提出模型进行对比验证本文模型的有效性,包括HTC数据集提出的基于SVM的基准模型^[1]、基于CFS的LightGBM模型^[9]、基于集成LightGBM与多层感知机的分类器模型^[17]、一维DenseNet模型^[11]、EmbraceNet模型^[12]以及双向LSTM模型^[21]。在比较各个方法的性能过程中,针对HTC数据集的原始数据格式,本文运用相同的处理方式以及划分方式进行对比,避免了数据处理过程中引入的人为误差影响实验过程的可靠性。

HTC数据集中包含的传感器类型有加速度、地磁与陀螺仪,现有工作中采用其他传感器的算法将会针对该数据集进行重新适配。HTC数据集中传感器的每个样本分别为形如 (x, y, z) 的传感器数据三元组,为了丰富传感器原始观测数据,本文使用式(6)~(9)对每个传感器再提取了4种分量,最终构成传感器数据七元组 (x, y, z, f, h, i, d) 。

$$f = \sqrt{x^2 + y^2 + z^2} \quad (6)$$

$$h = \sqrt{x^2 + y^2} \quad (7)$$

$$i = \arcsin(z/f) \quad (8)$$

$$d = \arcsin(y/h) \quad (9)$$

其中:式(6)计算了传感器向量的总强度 f ;式(7)计算了总强度 f 在水平面上的投影,即水平强度 h ;式(8)计算了总强度 d 与水平面的夹角,即倾角 i ;式(9)计算了水平强度偏离北向的角度,即偏角 d 。提取上述分量后能够进一步提取传感器的

物理特性,从而便于神经网络挖掘深层次特征。

用户乘坐交通模式下采集到的传感器数据在时间上具有连续性。在HTC交通模式识别数据集中,本文使用时长为5 s的时间窗口构建一个样本。本文对不同传感器数据在时间窗口内进行下采样来保证不同传感器数据长度一致。本文抽取了部分交通模式下某一时间窗口内传感器的总强度变化表现,如图5所示。

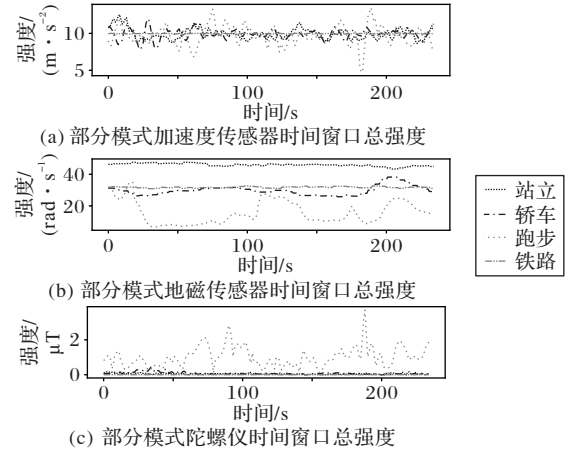


图5 部分交通模式的传感器总强度变化示意图

Fig. 5 Schematic diagram of total sensor intensity changes of some traffic modes

从传感器数值的总强度变化可见不同交通模式下具有明显不同信号变化趋势。图5(a)体现出用户在跑步过程中加速度总强度变化幅度较为明显,同样在图5(c)中用户在跑步过程中陀螺仪变化也较为明显。此外,在行驶状态变化较为明显的轿车交通模式中,如图5(a)所示加速度变化也较为明显。由上述例子可见在不同交通模式下,通过对传感器信号进行时序特征挖掘,使模型具有达到区分不同交通模式的潜力。

3.2 模型训练

在模型训练过程中,为了解决模型受传感器特征不同量纲的影响导致梯度收敛过慢甚至难以收敛的问题,本文对原始传感器数据进行了健壮归一化(Robust Normalization),对单一传感器数据 X 的健壮归一化过程如式(10)所示:

$$S_{\text{robust}} = \frac{\text{clip}(X, B_{\text{lower}}, B_{\text{upper}})}{B_{\text{upper}} - B_{\text{lower}}} \quad (10)$$

$$B_{\text{lower}}(X) = Q_1(X) - 1.5(Q_3(X) - Q_1(X)) \quad (11)$$

$$B_{\text{upper}}(X) = Q_3(X) + 1.5(Q_3(X) - Q_1(X)) \quad (12)$$

其中, Q_1 、 Q_3 分别代表上四分位数和下四分位数。健壮归一化能够确保归一化过程中传感器特征不受异常噪声的影响,将归一化范围确定在大多数正常信号范围内。

本文在训练残差时域注意力网络的过程中使用自适应均方根梯度(Adaptive Mean Square Gradient, AMSGrad)^[28]作为模型的优化器,起始学习率设为0.000 1。将训练数据按每批128个样本训练200轮。本文提出的模型采用4个残差单元堆叠构建的残差神经网络提取每个传感器128个通道的特征,此外,本文提出的模型利用3层膨胀卷积块堆叠的时域卷积网络对每个传感器提取256个通道的全局特征。

本文模型训练目的是不断优化模型参数,并降低模型输出的软标签与实际交通模式标签之间的误差,详见式(13), $\hat{y}$

为本文模型输出的预测结果,该结果与样本所属真实类别计算交叉熵作为模型损失函数。

$$L(\hat{y}, class) = -\hat{y}[class] + \ln\left(\sum_j e^{\hat{y}[j]}\right) \quad (13)$$

3.3 总体分类性能对比

在实验过程中,本文以准确率作为主要模型评价指标,同时引入每种交通模式下的召回率与准确率来进一步分析模型的识别性能。在使用 HTC 交通模式数据集进行实验过程中,本文选取已有的具有代表性的交通模式识别方法作为基准,其中准确率对比实验结果如表 1 所示。可以看出,本文提出的残差时域注意力网络能够在测试集上达到 96.07% 的准确率,而其他基准模型中只有文献[11]模型,文献[12]模型能够达到高于 90% 的准确率。本文提出的模型在 HTC 数据集下的识别准确率均高于其他基准算法。为了进一步深入分析各个交通模式的识别情况,本文对上述算法的召回率与精确率进行了对比,并分析不同交通模式在模型识别中的难易程度。

表 1 不同交通模式识别算法的识别准确率对比

Tab. 1 Recognition accuracy comparison of different traffic mode recognition algorithms

方法	准确率/%
文献[1]方法	75.36
文献[9]方法	82.68
文献[17]方法	86.21
文献[11]方法	91.51
文献[12]方法	90.15
文献[21]方法	85.76
本文方法	96.07

3.4 单一交通模式识别性能对比

由表 2 与表 3 可见,对所有基准算法和本文提出的模型分别在不同交通模式下的召回率与精确率进行了计算。召回率

与精确率计算方式如式(14)与式(15)所示:

$$recall(X_c) = \frac{TP(X_c)}{TP(X_c) + FN(X_c)} \quad (14)$$

$$precision(X_c) = \frac{TP(X_c)}{TP(X_c) + FP(X_c)} \quad (15)$$

其中: c 代表某一类交通模式,而 X_c 指所有交通模式 c 下的样本; TP (True Positive)、 FP (False Positive)、 FN (False Negative)与 TN (True Negative)分别指在样本真实情况正例预测结果为正例、真实情况负例预测结果为正例、真实情况正例预测结果为负例,以及真实情况负例预测结果为负例的个数。

召回率可以计算出在交通模式 c 下,模型能够在所有该类别样本中正确识别该交通模式的比例,而精确率可以计算出分类方法在预测结果中能够准确预测交通模式 c 的比例。从表 2 中的数据可分析出,在所有交通模式识别算法中摩托车与公交车两种交通模式与其他交通模式相比在测试集中的召回率较差。由表 3 可见,摩托车与自行车两种交通模式在查准结果中能够符合真实交通模式的比例普遍较低,而站立、行走以及跑步这类与用户姿态变化相关的交通模式往往更容易识别。其中文献[1]中 SVM 分类器能够在公交车交通模式上达到较高的精确率,但该模型的召回率过低,上述情况也表明了公交车交通模式在该算法上难以在实际情况中得到准确的识别,同理这个现象也出现在其他交通模式识别基准模型中,但要优于 SVM 分类器。本文提出的方法在识别公交车交通模式时相较基准模型在各个指标上都有所提升,但召回率也低于精确率,这也体现出准确识别公交车交通模式具有一定的挑战,传统方法将实际的公交车模式大量误判为其他交通模式,仅仅能够区分出部分情况下的公交车模式。此外,本文提出的模型能够在摩托车、自行车等在其他基准模型中较难区分的交通模式中得到较高精确率与召回率。

表 2 不同识别算法的识别召回率对比

Tab. 2 Recognition recall comparison of different classification algorithms

方法	召回率/%							
	站立	行走	跑步	自行车	摩托车	轿车	公交车	轨道
文献[1]方法	76.60	73.52	72.98	69.98	64.87	78.33	21.27	85.96
文献[9]方法	88.09	80.61	96.01	80.56	72.20	83.85	70.34	85.90
文献[17]方法	91.31	85.56	96.50	82.10	78.65	89.28	66.60	88.44
文献[12]方法	93.28	91.18	97.08	87.93	89.14	90.17	71.30	91.39
文献[11]方法	95.72	92.31	97.40	93.64	89.40	90.55	74.96	92.16
文献[21]方法	92.84	85.59	95.98	77.92	83.03	85.81	65.20	88.04
本文方法	97.69	96.28	97.75	94.91	94.83	96.36	90.34	96.73

表 3 不同识别算法的识别精确率对比

Tab. 3 Recognition precision comparison of different classification algorithms

方法	精确率/%							
	站立	行走	跑步	自行车	摩托车	轿车	公交车	轨道
文献[1]方法	84.67	69.53	95.62	72.39	72.25	74.60	97.46	74.05
文献[9]方法	96.28	74.79	95.21	74.16	72.50	84.74	84.14	84.68
文献[17]方法	93.19	79.79	91.86	82.91	80.92	85.38	91.35	88.53
文献[12]方法	93.49	92.29	98.49	87.72	85.33	90.41	81.51	90.78
文献[11]方法	93.72	92.46	98.36	90.73	89.37	90.77	80.16	92.58
文献[21]方法	92.95	85.47	96.79	80.17	80.30	85.18	70.29	88.15
本文方法	97.43	95.74	99.04	93.93	94.36	96.51	94.06	96.64

综合表 2~3 可以分析出,本文提出的模型在多种交通模式下召回率与精确率相差较小,也表明本文模型在识别其他

交通模式下具有较高鲁棒性。

本文还在 HTC 数据集下对本文提出的模型计算了混淆

矩阵,如表 4 所示。从表 4 中可以看出,本文提出的模型将部分属于公交车交通模式类别的样本识别为轨道交通模式和轿车交通模式。这几种交通模式具有一定的相似性且用户在这几种交通模式下运动姿态变化相对较小,运动状态较为稳定,当用户姿态变化较大时比如跑步、行走等模式下模型能够达到较好识别效果。

表 4 残差时域注意力模型识别效果混淆矩阵

Tab. 4 Confusion matrix of recognition results by residual temporal attention network

单位: %
unit: %

预测类别	真实类别							
	站立	行走	跑步	自行车	摩托车	轿车	公交车	轨道
站立	96.49	0.25	0.46	0.36	0.20	0.14	0.28	0.28
行走	0.39	95.63	0.08	0.75	1.42	0.26	0.43	0.28
跑步	0.01	0.31	98.90	0.11	0.31	0.02	0.00	0.01
自行车	0.37	0.57	0.11	94.65	1.77	0.12	0.48	0.27
摩托车	0.45	1.18	0.31	1.94	93.25	0.99	1.23	0.56
轿车	0.71	0.57	0.00	0.61	1.48	96.16	1.54	1.30
公交车	0.14	0.21	0.00	0.33	0.39	0.61	93.84	0.75
轨道	1.39	1.24	0.11	1.21	1.14	1.67	2.17	96.52

3.5 消融实验

本文提出的残差时域注意力模型在交通模式识别过程中分别针对不同时间尺度下进行了时序特征挖掘,该模型分别使用残差网络与时域卷积网络对短期时序特征与长期时序特征进行特征提取。本文首先通过消融实验对本文提出的交通模式识别模型中上述两个模块的影响进行了评估,其次本文提出的模型对残差网络与时域卷积网络提取的特征进行基于通道的注意力机制特征校准进行了分析。基于上述模型消融实验设计所得到的实验结果如表 5 所示。

表 5 本文模型消融实验准确率对比

Tab. 5 Accuracy comparison of proposed model in ablation experiments

组合	组合方式	准确率/%
1	去除残差网络	91.39
2	去除时域卷积网络	80.27
3	去除残差网络与时域卷积网络	65.68
4	去除通道注意力机制	85.77
5	去除通道注意力机制与残差网络	59.58
6	去除通道注意力机制与时域卷积网络	78.21
7	最终模型	96.07

由表 5 可见,本文提出的模型在缺少残差网络提取的短期时序特征时,模型在识别准确率上下降较少,但当模型缺少时域卷积网络提取的长期时序特征时,模型在识别准确率上下降较大。这也说明了全局特征在本文模型中对交通模式识别起到了更为重要的影响。当本文模型只利用传统的卷积网络进行特征提取时,该模型在测试集上的识别精度下降明显,只有 65.68% 的识别准确率,这表明残差网络 and 时域卷积网络在本文模型中对交通模式识别所提取的特征是十分有效的。本文模型包含针对传感器的残差网络、通道注意力机制以及针对特征空间的时域卷积网络,而当本文模型移除了基于注意力机制的特征重校准之后,模型精度从 96.07% 下降到 85.77%,这表明模型的特征重校准模块有效地筛选出了交通模式识别中的重要特征,并帮助提高了模型识别的鲁棒性。从表 5 中可以发现,本文所提出模型中各个模块均会对交通模式识别起到重要帮助作用。

由此可见,残差网络提供的细粒度感知能力在模型中起到重要作用。组合 7 相较于组合 1 提高了 4.68 个百分点,而

相较于组合 5 可以对比得出,综合通道注意力机制与残差网络的表现,仅仅使用时域卷积网络的组合 5 效果最差。通道注意力机制与残差网络均能够使整个模型对多源传感器的特征通道再平衡,但残差网络难以对不同通道特征进行通道级别的重新校准,只能通过卷积操作的细粒度权重进行简单平衡。只有通道注意力机制与残差网络配合起来分别在特征尺度与通道尺度进行重校准与特征提取,模型性能能够得到较大的提升,而任一部分缺失都能够极大影响模型总体的性能。其次,由组合 2 也可以得出,缺失粗粒度时域特征感知能力的模型效果也难以达到理想水平,时域卷积网络在模型中起到了至关重要的作用。

3.6 算法计算开销对比

本文针对上述不同分类器算法进行了运算时间开销与空间占用的分析与对比。其中,针对基于传统机器学习算法的分类器方法,本文将针对原始数据中验证集部分进行特征工程处理的空间与时间开销也计入算法的时间与空间开销中。对于不同方法,本文通过统计分类器方法在验证集部分数据上花费的时间作为时间开销的指标,通过统计在验证集部分数据上花费的最大内存与显存占用值作为内存占用的指标,从而对不同方法进行针对计算开销的比较,详细计算开销对比如表 6 所示。

表 6 不同交通模式识别算法的计算开销对比

Tab. 6 Comparison of computational overhead for different traffic mode recognition algorithms

方法	验证集预测 时间开销/s	验证集预测 占用空间/MB
文献[1]方法	1 278.677	191.260
文献[9]方法	2.488	151.939
文献[17]方法	20.507	298.844
文献[11]方法	28.697	1.946
文献[12]方法	26.043	350.788
文献[21]方法	25.983	7.091
本文方法	32.146	19.908

由计算开销对比可以看出,本文算法在时间开销与空间开销上能够达到相对较低的占用。由本文提出算法的时间与空间开销测试结果可见,在计算开销与算法精度上本文算法能够达到较好的平衡。

4 结语

本文基于加速度、地磁以及陀螺仪等低功耗传感器数据提出了残差时域注意力模型进行用户所处交通模式的识别。该模型通过引入残差网络与时域卷积网络进行不同感受野的时序特征挖掘,并使用基于通道的注意力机制对传感器特征进行特征重校准,使得该模型在不同交通模式下获得更高的检测鲁棒性。本文所提算法在 HTC 交通模式数据集上取得了良好效果,达到了 96.07% 的准确率,远高于其他现有基准算法。但本文算法仍有提升识别准确度的空间,同时随着移动智能终端的发展成熟,移动设备也具有适配该算法的潜力。未来将克服目前局限,进一步提升该交通模式识别算法性能,扩展交通模式识别算法新应用。

参考文献 (References)

- [1] YU M C, YU T, WANG S C, et al. Big data small footprint: the design of a low-power classifier for detecting transportation modes [J]. *Proceedings of the VLDB Endowment*, 2014, 7(13): 1429-1440.
- [2] WANG L, GJORESKE H, CILIBERTO M, et al. Enabling reproducible research in sensor-based transportation mode recognition with the Sussex-Huawei dataset [J]. *IEEE Access*, 2019, 7: 10870-10891.
- [3] ZHENG Y, XIE X, MA W. GeoLife: a collaborative social networking service among user, location and trajectory [EB/OL]. [2020-03-30]. <http://sites.computer.org/debull/A10june/geolife.pdf>.
- [4] YANG H F, DILLON T S, CHEN Y P P. Optimized structure of the traffic flow forecasting model with a deep learning approach [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(10): 2371-2381.
- [5] LV Y, DUAN Y, KANG W, et al. Traffic flow prediction with big data: a deep learning approach [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(2): 865-873.
- [6] 汪磊, 左忠义, 傅军豪. 基于 SVM 的出行方式特征分析和识别研究 [J]. *交通运输系统工程与信息*, 2014, 14(3): 70-75, 84. (WANG L, ZUO Z Y, FU J H. Travel mode character analysis and recognition based on SVM [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2014, 14(3): 70-75, 84.)
- [7] ZHENG Y, ZHANG L, XIE X, et al. Mining correlation between locations using human location history [C]// *Proceedings of the 2009 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. New York: ACM, 2009: 472-475.
- [8] GHOSE A, LI B, LIU S. Mobile targeting using customer trajectory patterns [J]. *Management Science*, 2019, 65(11): 4951-5448.
- [9] 熊苏生. 基于改进 LightGBM 的交通模式识别算法 [J]. *计算机与现代化*, 2018(10): 68-73, 126. (XIONG S S. Identifying transportation mode based on improved LightGBM algorithm [J]. *Computer and Modernization*, 2018(10): 68-73, 126.)
- [10] ZHENG L, LI S, ZHU C, et al. Application of IndRNN for human activity recognition: the Sussex-Huawei locomotion-transportation challenge [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 869-872.
- [11] ZHU Y, ZHAO F, CHEN R. Applying 1D sensor DenseNet to Sussex-Huawei locomotion-transportation recognition challenge [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 873-877.
- [12] CHOI J H, LEE J S. EmbraceNet for activity: a deep multimodal fusion architecture for activity recognition [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 693-698.
- [13] BALABKA D. Semi-supervised learning for human activity recognition using adversarial autoencoders [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 685-688.
- [14] FRIEDRICH B, CAUCHI B, HEIN A, et al. Transportation mode classification from smartphone sensors via a long-short-term-memory network [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 709-713.
- [15] CARPINETI C, LOMONACO V, BEDOGNI L, et al. Custom dual transportation mode detection by smartphone devices exploiting sensor diversity [C]// *Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications Workshops*. Piscataway: IEEE, 2018: 367-372.
- [16] WIDHALM P, LEODOLTER M, BRÄNDLE N. Ensemble-based domain adaptation for transport mode recognition with mobile sensors [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 857-861.
- [17] LU H, PINAROC M, LV M, et al. Locomotion recognition using XGBoost and neural network ensemble [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 757-760.
- [18] AHMED M, ANTAR A D, HOSSAIN T, et al. POIDEN: position and orientation independent deep ensemble network for the classification of locomotion and transportation modes [C]// *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers*. New York: ACM, 2019: 674-679.
- [19] COROAMĂ V C, TÜRKE C, MATTERN F. Exploring the usefulness of Bluetooth and WiFi proximity for transportation mode

- recognition [C]// Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers. New York: ACM, 2019: 37-40.
- [20] STENNETH L, WOLFSON O, YU P S, et al. Transportation mode detection using mobile phones and GIS information [C]// Proceedings of the 2011 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: ACM, 2011: 54-63.
- [21] ALWAN A, FREY V, LAN G L. Orange labs contribution to the Sussex-Huawei locomotion-transportation recognition challenge [C]// Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing/ 2019 ACM International Symposium on Wearable Computers. New York: ACM, 2019: 680-684.
- [22] SU X, CACERES H, TONG H, et al. Online travel mode identification using smartphones with battery saving considerations [J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(10): 2921-2934.
- [23] GJORESKE H, CILIBERTO M, WANG L, et al. The university of Sussex-Huawei locomotion and transportation dataset for multimodal analytics with mobile devices [J]. IEEE Access, 2018, 6: 42592-42604.
- [24] REDDY S, MUN M, BURKE J, et al. Using mobile phones to determine transportation modes [J]. ACM Transactions on Sensor Networks, 2010, 6(2): No. 13.
- [25] HUANG G, LIU Z, MAATEN L V D, et al. Densely connected convolutional networks [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 2261-2269.
- [26] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [EB/OL]. [2021-03-12]. <http://arxiv.org/pdf/1803.01271.pdf>.
- [27] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [28] REDDI S J, KALE S, KUMAR S. On the convergence of ADAM and beyond [EB/OL]. [2021-03-12]. http://www.sanjivk.com/AdamConvergence_ICLR.pdf.

This work is partially supported by the National Key Research and Development Program of China (2019YFC1511400), the National Natural Science Foundation of China (61872046), the Beijing Natural Science Foundation (4212024), the Scientific and Technological Innovation Ability Improvement Action Program of Beijing University of Posts and Telecommunications (2019XD-A06), the Joint Research Fund of Beijing Natural Science Foundation and Haidian Original Innovation (L192004), the Key Research and Development Project of Hebei Province (19210404D), the Key Technology Research Program of Inner Mongolia Autonomous Region (2019GG328), the Open Project of the Beijing Key Laboratory of Mobile Computing and Pervasive Device.

LIU Shize, born in 1988, M. S. His research interests include big data mining, intelligent perception.

ZHU Yida, born in 1996, Ph. D. candidate. His research interests include urban computing, transportation mode recognition.

CHEN Runze, born in 1996, Ph. D. candidate. His research interests include urban computing, transportation mode recognition, location semantic mining.

LUO Haiyong, born in 1967, Ph. D., associate research fellow. His research interests include mobile intelligence, pervasive computing.

ZHAO Fang, born in 1968, Ph. D., professor. Her research interests include mobile internet, big data mining.

SUN Yi, born in 1979, M. S., senior engineer. His research interests include mobile internet, big data mining.

WANG Baohui, born in 1973, M. S., professorate senior engineer. His research interests include software engineering.