

基于多尺度特征提取的交通模式识别算法

刘世泽¹, 秦艳君², 王晨星², 高存远², 罗海勇^{3*}, 赵方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, 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:针对普适交通模式的场景感知功耗高、场景复杂的问题,提出一种融合残差网络(ResNet)和带孔卷积的交通模式识别算法。首先,使用快速傅里叶变换(FFT)将一维传感器数据转换为二维频谱图像;然后,使用主成分分析(PCA)算法对频谱图像降采样;最后,使用ResNet挖掘交通模式的局部特征,使用带孔卷积挖掘交通模式的全局特征,从而实现对八种交通模式进行识别。与决策树、随机森林、AlexNet等八种算法在实验中的对比评估结果显示,融合ResNet和带孔卷积的交通模式识别算法在静止、走路、跑步等八类交通模式上均有最高准确率。该算法具有良好识别精度和鲁棒性。

关键词:行为识别;交通模式识别;残差网络;带孔卷积;低功耗

中图分类号:TP182 **文献标志码:**A

Transportation mode recognition algorithm based on multi-scale feature extraction

LIU Shize¹, QIN Yanjun², WANG Chenxing², GAO Cunyuan², LUO Haiyong^{3*},

ZHAO Fang², WANG Baohui¹

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

2. School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China;

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

Abstract: Aiming at the problems of high power consumption and complex scene for scene perception in universal transportation modes, a new transportation mode detection algorithm combining Residual Network (ResNet) and dilated convolution was proposed. Firstly, the 1D sensor data was converted into the 2D spectral image by using Fast Fourier Transform (FFT). Then, the Principal Component Analysis (PCA) algorithm was used to realize the downsampling of the spectral image. Finally, the ResNet was used to mine the local features of transportation modes, and the global features of transportation modes were mined with dilated convolution, so as to detect eight transportation modes. Experimental evaluation results show that, compared with 8 algorithms including decision tree, random forest and AlexNet, the transportation mode recognition algorithm combining ResNet and dilated convolution has the highest accuracy in eight traffic patterns including static, walking and running, and the proposed algorithm has good identification accuracy and robustness.

Key words: activity recognition; transportation mode recognition; Residual Network (ResNet); dilated convolution; low power consumption

0 引言

在移动互联网时代,智能手机的功能越来越强大,使得智能手机移动设备的传感功能的可编程性得到了广泛的应用,使其拥有很高的市场普及率,再加上第三方应用软件的支持和智能手机快速的发展,就成为了监控人们行为的有效工具。因此基于移动智能终端的场景感知成为场景感知的主要方式。

交通模式识别被认为是一种特殊的场景感知^[1-2]。通过

智能移动设备捕获个人交通行为,这对许多领域有着非常积极的影响。例如:对个人交通行为的智能监控,可以实现对移动跟踪,也可以利用这些信息改善城市规划、提供合适的道路出行建议,快速为救护车安排最合适的行驶方案,以及追踪用户碳足迹,呼吁人们低碳出行。最后,智能监测可以作为用户行为分析的一部分,进行实时旅程规划或针对性的广告投放。最广泛的电子健康、营销、智能城市、交通管理和安全,其中有关参与者流动性的信息构成了所收集数据上下文的一个重要维度。

收稿日期:2020-11-04;修回日期:2021-04-02;录用日期:2021-04-12。

基金项目:国家自然科学基金资助项目(61872046);北京邮电大学提升科技创新能力行动计划项目(2019XD-A06)。

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; 秦艳君(1991—),女,山西长治人,博士研究生,CCF会员,主要研究方向:城市计算、交通模式识别; 王晨星(1997—),男,河北衡水人,硕士研究生,主要研究方向:城市计算、交通模式识别、交通流预测; 高存远(2000—),男,云南昆明人,主要研究方向:城市计算、交通模式识别; 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; 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。

目前有研究利用全球定位系统(Global Positioning System, GPS)的数据来检测运输方式^[3-4]。然而,基于GPS的方法存在以下缺点:1)GPS信号在城市峡谷、隧道或屏蔽区域,会出现信号丢失和错误。2)GPS传感器消耗大量的能量,并可能迅速耗尽移动设备的电池;融合GPS和地理信息系统(Geographic Information System, GIS)的交通模式识别算法,该方法利用运输网络信息,包括实时总线的位置、空间轨道和巴士站信息,改善了GPS的数据对交通工具识别判断不准的一些情况,但是现在的GIS信息系统更新得仍旧不够实时,存在功耗高的问题。

大量研究使用基于机器学习的算法实现交通模式识别,例如AdaBoost分类器、支持向量机(Support Vector Machine, SVM)^[5-6]、决策树、随机森林^[7-8]、贝叶斯网络^[9-10]、分层的分类器^[11-14]等算法识别交通工具模式。但是这些算法需要首先根据人们的先验知识完成特征提取,而实现人工特征提取过程存在复杂、耗时和识别准确率低等问题。目前基于加速度传感器的交通模式识别方法提取帧特征、段特征和峰特征,并使用这些特征进行交通模式识别,但是该方法易受重力的影响,也很容易受到采集设备的噪声干扰,这样就会使收集数据时产生数据不准确,导致最终判断准确率低。

针对上述问题,研究者们提出了基于深度学习的交通模式识别算法,如:常见深度学习算法有基于循环神经网络(Recurrent Neural Network, RNN)、卷积神经网络(Convolutional Neural Network, CNN)^[15]、双向长短期记忆网络(Bidirectional Long Short-Term Memory, BiLSTM)^[16]和长短期记忆(Long Short-Term Memory, LSTM)网络^[17-18]等算法。但这些算法存在功耗高、训练时间长和精度低等问题。本文提出了一种融合残差网络和带孔卷积的算法,实现交通模式识别,主要工作如下:

1)使用低功耗智能手机低功耗传感器,摒弃GPS、Wi-Fi等功耗高的传感器。使用低功耗的传感器包含加速度传感器、气压传感器、地磁、陀螺仪。将这些传感器得到的原始数据,通过快速傅里叶变换(Fast Fourier Transform, FFT)将时序数据转换为频域数据(频谱图)。将时序数据转换成图片数据的做法,可以有效消除采集原始数据从不同部位和不同的从而获取更稳定特征。

2)使用残差网络(Residual Network, ResNet)学习交通模式短时间的局部特征。由于神经网络的每一层对应提取特征不同层次特征学习,而且网络结构越深,提取的不同层次的特征越多,不同层次的信息组合也越多。由于网络结构越深,会导致网络的梯度消失、梯度爆炸和网络性能退化问题,造成错误率上升。使用残差网络可以解决上述这些问题,更好地提升网络性能,从而达到深度提取各类交通模式的深层局部特征的目的。

3)利用带孔卷积学习长时间的全局特征。对于特征提取不仅仅需要从局部提取特征,更需要从全局提取特征,去获得更好的识别准确率。在不损失数据信息的情况下增大感受野,能有效地提取基于各种交通工具的全局特征;并且可以有效降低信息的复杂度,减少运算复杂度。

1 相关工作

在目前交通模式识别领域,已经有很多的学者在研究该

问题,从算法层面主要分为两类:一类是利用传统学习方法,主要依赖手工提取特征,实现交通模式识别;另一类是利用神经网络方法,通过神经网络实现各类交通模式的深度特征提取,实现交通模式识别。

文献[19]提出了一种利用智能手机位置信息和各种可用的地理信息(如公交站点位置、铁路线路和公交位置信息)实时检测用户出行方式的方法。与之前的相关工作不同,用于定位的采样率为30 s。对于接近实时的分类,作者使用了一个1 min的移动记录窗口(相当于两个位置记录),它使用了以下属性:平均速度、平均加速度和的变化。该方法利用的是传统的手工提取特征的方法。

文献[20]提出了一个基于Bi-LSTM模型的自动轨迹分类框架,以将原始轨迹分类为不同的运输方式。通过嵌入时间间隔作为外部特征进一步提高了性能。该框架不需要耗时的复杂的人工特征提取过程,而是通过Bi-LSTM框架学习特征。

文献[21]利用基于深度神经网络(Deep Neural Network, DNN)的方法,将多层神经元堆叠起来优化权重,从而进行交通模式识别。

文献[22]提出了一种基于树的集成分类器,通过原始数据,手工提取全局特征和局部特征相结合,因此需要一个复杂和费时的特征工程过程。

在文献[23]中基于CNN的架构,利用原始的GPS轨迹来预测出行方式,创建了一个四通道输入量,其中每个通道代表一个基本运动学属性序列,包括速度、加速/减速、跳动和承载率,实现对各类交通模式进行区分。

综上所述,人工提取特征存在着一些重要缺陷,主要包括对交通和环境条件的脆弱性,以及在提取高效特征时,对存在手机型号和采用用户不具有普适性。为解决这些问题考虑利用神经网络代替传统方法,从而自动地从原始输入驱动高级特征;但是现在单一的一些模型并不能很好地处理原始的时间序列数据。

2 融合残差网络和带孔卷积框架

本文提出了融合残差网络和带孔卷积的算法,实现交通模式识别。本文算法总体框架如图1所示,主要包含三个部分,分别是:数据预处理、各类交通工具局部特征提取和各类交通工具全局特征提取。这三个部分具体内容如下:

1)数据预处理。使用手机采集原始数据,采用不同手机采集数据,会产生不同设备之间采集误差。采集者会放置不同的部位,例如:手持、背包、手臂和裤袋等姿态,对数据采集产生不同姿态误差影响。为了避免这些误差的影响,利用手机传感器采集原始数据,通过FFT算法,将时间序列的数据转换成图片数据,可以有效地去除上述误差影响。

2)局部特征。由于ResNet网络结构,对于图片数据,通过深度的学习图片数据,可以有效地提取交通工具深层局部特征。

3)全局特征提取。带孔卷积可以在不损失图片数据信息的情况下,增大感受野。通过每个增大卷积的卷积核,可以扩大感受野,获得长时间依赖关系,且没有增加参数数量,从而有效地捕获八类交通模式图片数据的全局特征。

最后将全局特征和局部特征通过四层全联接层,最后通过SoftMax得到八类交通模式分类结果。

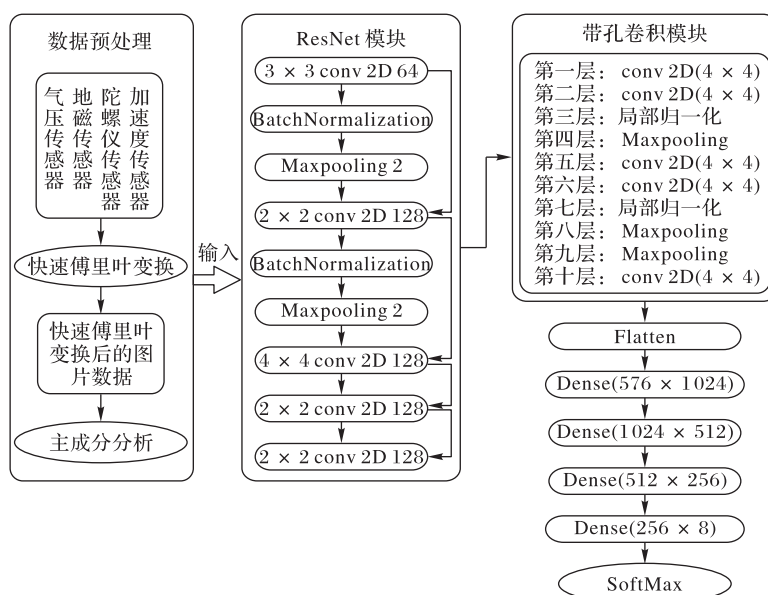


图1 融合残差网络和带孔卷积交通模式识别算法框架

Fig. 1 Framework of transportation mode detection algorithm combining residual network and dilated convolution

3 数据采集和数据预处理

3.1 数据采集

利用手机采集与交通模式相关数据,进行交通模式识别,已经是非常普遍的情况^[20-21]。

在本文中,数据采集人员把收集的数据和采集人员的日常生活中想要做的事情结合起来,例如:参观博物馆,在一个特定的餐馆吃午饭,或者骑自行车去拜访朋友等。这样数据集可以包含真实的活动和典型位置。

数据收集使用了四款华为 Mate 9 手机,放在人们通常携带手机的地方:第一部手机,采集者会拿在手里,或者放在胳膊上戴的臂带中,在数据集中,这个位置称为“手(Hand)”；第二部手机,采集者放在胸袋或胸带上(例如跑步时),在数据集中,此位置称为“躯干(Torso)”；第三部手机,采集者放在臀部的高度,通常放在裤子的前口袋里,在数据集中,此位置称为“裤袋(Hips)”；第四部手机,采集者放在背包或侧包里,在数据集中,这个位置称为“包(Bag)”。这些手机配备了专门采集数据的数据采集器。

3.2 数据预处理

本文采用了由英国萨塞克斯大学(University of Sussex)和华为公司发布的用于移动用户多模态运动分析和交通模式识别数据集(Sussex-Huawei Locomotion dataset, SHL)。

SHL 数据集用途广泛,主要可以用于:用户行为识别、多模态无线网络覆盖范围估计、卫星覆盖范围估计、路况分析、交通状况分析、谷歌公司行为识别应用程序接口(Application Programming Interface, API)、活动发现、传感器融合定位、基于视觉和声音的行为识别等方面。

根据本文对数据集的要求,本文对数据集做了一定处理。其中,每个传感器数据文件中包含 196 072 行 $\times$ 500 列,对应于 196 072 帧,每个帧包含 500 个样本(采样频率为 100 Hz)。

每一个样本通过快速傅里叶变换,将其转化为一张图片。使用滑动窗口大小是 400,滑动的步长是 1。

如图 2 所示,就是将加速度 x 轴的原始数据,通过快速傅

里叶变换方法,将原始时间序列数据转换为图片频域数据。如图 3 所示,是在不同位置的数据采集效果呈现,可以看出不同的位置,加速度 x 轴的数据变化量很大。从图 3 可以看出,时间序列的数据比较分布不均匀。在一定程度上不同采集者和位置都会有一定的变化,但是通过 FFT 转换成图片(图 4 所示)数据波动就会很小。

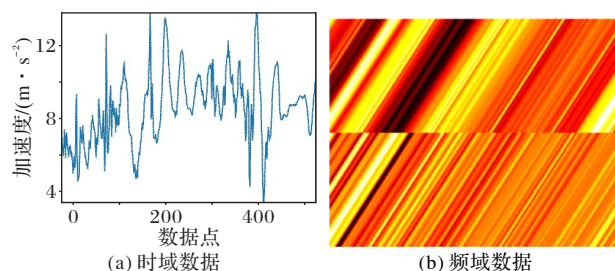


图2 将时域数据转换为频域图片数据过程

Fig. 2 Process of converting time domain data into frequency domain image data

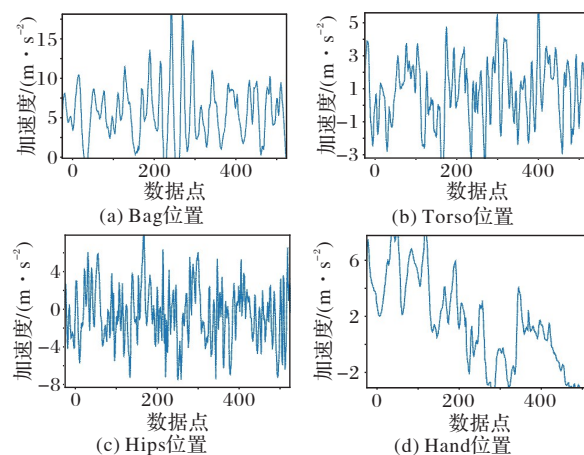


图3 不同位置不同采集者采集的原始数据

Fig. 3 Raw data collected by different users in different locations

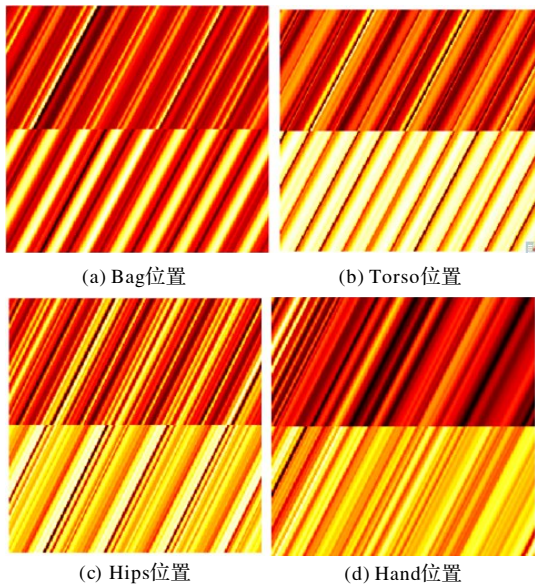


图4 快速傅里叶变换后的图片数据

Fig. 4 Image data after FFT

3.3 数据主成分分析

经过FFT处理后的表示各种交通模式特征图片,存在图片过大的问题。因此会导致出现模型训练复杂度过大问题,故需要对图片进行降采样。如果使用普通的插值算法,可能会导致图片细节过多地丢失,从而降低模型的精度。如算法1所示,本文使用主成分分析(Principal Component Analysis, PCA)方法是先将图片二维矩阵展开为向量,并进行主成分分析,选取所有位置中特征值最大的50 176个,将这些像素再重新组合为新的图片,大小为 224×224 。

算法1 数据PCA预处理。

输入 原始图像数据 $data$ 。

```

 $dataD \leftarrow$  去中心化( $data$ );
 $dataF \leftarrow$  Flatten( $dataD$ );
 $dataC \leftarrow$  计算协方差矩阵( $dataF$ );
 $dataS \leftarrow$  SVD分解( $dataC$ );
 $dataK \leftarrow$  sort( $dataS$ )[0:50 176];
 $dataR \leftarrow$  dot( $dataD$ ,  $dataK$ );
 $dataE \leftarrow$  reshape( $dataR$ )

```

首先,对数据进行去中心化并展平;接下来,计算整个矩阵的协方差矩阵,对协方差矩阵进行奇异值分解(Singular Value Decomposition, SVD)选取特征值最高的50 176个特征向量再组成矩阵;最后,将这个矩阵与原来的矩阵相乘,并转换回大小为 224×224 的图片的形式。此时保留下来的特征点拥有着较大的方差,根据最大方差理论,这些特征点保留了原始图像中的大部分关键数据。

4 局部特征提取

卷积神经网络(CNN),能够有效地提取浅层/中级/深层特征。一般来说,层数越多,提取到的不同层次的特征就会更丰富、抽象以及富含语义信息。但是要想实现这个目标,往往需要加深CNN网络层为代价,同时常常带来参数量大、梯度消失、梯度爆炸等问题,导致网络无法学习到准确特征。

ResNet的主要思想是,假定某部分神经网络的输入为 x ,期望的输出为 $H(x)$ 。 $H(x)$ 是期望的复杂潜在映射,如果进行学习,则训练难度较大。可以知道的是,当学习到较为饱和

的准确率时或下层网络的误差开始增大时,接下来学习的目标就需要转变为恒等映射的学习,即:使输入 x 近似于输出 $H(x)$,以保持后面的网络层中不会造成精度继续下降。ResNet引入了捷径连接(Shortcut Connections)来试图解决这一问题。如图5所示,直接将输入 x 传到输出位置作为初始的结果,那么输出则变为 $H(x) = F(x) + x$ 。当 $F(x) = 0$ 时, $H(x) = x$,即恒等映射。因此,ResNet的学习目标则转变为残差 $F(x) = H(x) - x$ 。之后的网络层的学习目标全部转变为使残差 $F(x) \rightarrow 0$,这样就使得网络随着层数加深,精确率不会下降。

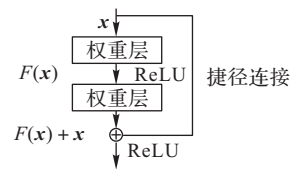


图5 ResNet单元结构

Fig. 5 Architecture of ResNet unit

残差网络可以有效地避免上述问题,同时学习准确各类交通模式特征。如图6所示,输入层是各类交通模式图片数据,数据大小是 (m, n, q) ,其中 m, n 是像素, q 是通道数。本文中的残差单元如式(1)、(2):

$$y_l + h(x_l) = \mathcal{F}(x_l, W_l) \quad (1)$$

$$x_{l+1} = f(y_l) \quad (2)$$

其中: x_l 是残差单元的输入(数据图片); x_{l+1} 表示残差单元的输出; f 表示激活函数(ReLU函数); $\mathcal{F}$ 是残差函数; $h(x_l) = x_l$; y_l 是第 l 层网络的输出; W_l 是可训练参数矩阵。

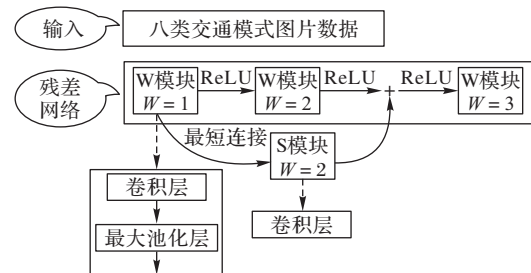


图6 局部特征提取网络结果

Fig. 6 Result of local feature extraction network

图6的中间模块是残差网络模块,该模块可以有效地提取交通模式的深度特征,W模块中,使用卷积网络实现特征提取过程。所以第 i 层残差单元,计算式如式(3)~(6):

$$w_1 = mp_{u_1}^{1, v_1} \{ f[conv_{u_1}^{1, v_1, p_1}(x_{l-1})] \} \quad (3)$$

$$w_2 = mp_{u_2}^{1, v_2} \{ f[conv_{u_2}^{1, v_2, p_2}(w_1)] \} \quad (4)$$

$$w_3 = mp_{u_3}^{1, v_3} \{ f(w_2 + s) \} \quad (5)$$

$$s = conv_{u_s}^{1, v_s, p_s}(x_l) \quad (6)$$

其中: w_i 是第 i 个W模块的输出, s 是S模块的输出, $mp_{u_i}^{l, v_i}(x_{u_i}^{l-1})$ 是最大池化操作; $conv_{u_i}^{l, v_i, p_i}(x_{u_i}^{l-1})$ 是卷积操作; v 是特征矩阵, u 是第 l 层卷积的单元。在本文设计的网络架构中,有三层残差网络模块。

通过每一层的残差单元,都能够提取到一定程度特征。通过加深残差网络的深度,就能够有效地提取八类交通模式

的深层特征。由于增加捷径连接,加入S模块,可以有效避免深度卷积网络带来的梯度消失和梯度爆炸问题。

对于每一层模块的超参数如表1所示。

表1 残差网络详细超参数

Tab. 1 Detailed hyperparameters in residual network

名称	值
残差卷积的滤波器(宽度,高度,通道数,输出大小)	[144, 144, 144, 100, 64, 8]
残差网络卷积大小	[4, 4]
残差正则化参数	0.000 5
残差卷积步长	[1, 1, 1, 1, 1]
残差网络最大池化大小	[2, 2]

5 全局特征提取

带孔卷积的作用是在同层计算不增加参数量的同时扩大感受野。感受野的大小决定了卷积过程中,一次获取信息的多少。如图7所示,对于输入,每一次映射的区域不同,第一层是每两个节点映射一个区域;第二层映射的区域面积变成每隔一个节点映射一次;第三层映射区域最大,变成每隔两个节点映射一次。

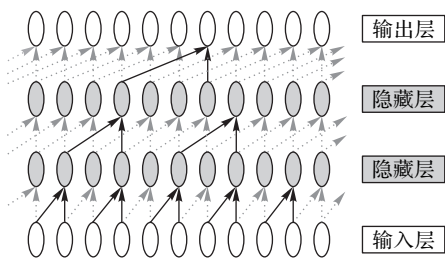


图7 空洞卷积表示

Fig. 7 Representation of dilated convolution

这样设计的好处是,能够在很大程度上减少参数量,且该网络能够有效保证保留内部数据结构(传统下采样方式就会失去数据内部结构,导致特征提取不够完整)。利用池化方法在增加感受野时,会将一些数据信息丢失,但是带孔卷积网络可以在不损失信息特征的情况下,增大感受野,获取更多信息。

最重要的是,带孔卷积网络能够准确获取多尺度信息特征,避免提取冗余特征,可以更好地挖掘各类交通模式数据之间的内部特征。本文利用带孔卷积从大尺度上提取全局特征,这些全局特征能够有效地增加交通模式识别的准确率。

表2是本文中带孔卷积设计到的超参数的详细描述,带孔卷积的卷积核计算不同于传统卷积计算方式。逐层增加了卷积核单元间隔距离,分别为0、1、2,在增加特征图感受野面积的同时有效控制了参数量的增加。

表2 带孔卷积网络详细超参数

Tab. 2 Detailed hyperparameters in dilated convolutional network

名称	值
带孔卷积网络卷积大小	[4, 4]
带孔卷积正则化参数	0.000 5
带孔卷积步长	[1, 1, 1, 1, 1]
残差网络最大池化大小	[2, 2]

本文算法的伪代码如算法2所示,主要分为三个阶段:数据输入、训练阶段和测试阶段。数据输入阶段主要将数据进

行了数据标准化、FFT,以及图像主成分分析等数据预处理方法。在训练和测试阶段,通过残差网络实现数据局部特征提取,带孔卷积实现数据的全局特征。然后,通过Adam优化器目标的训练权重收敛或训练达到预定轮数。最终,通过SoftMax实现八类交通模式识别。

算法2 融合残差网络和带孔卷积网络。

输入 SHL数据集传感器数据 *data*

初始化当前最佳精度为0

dataN ← Z-Score 标准化(*data*);

dataW ← 小波变换(*dataN*);

dataI ← 滑动窗口生成频谱图(*dataW*);

dataI2 ← 进行图像PCA(*dataI*);

dataTrain, *dataValidate* ← 随机分割(*dataI2*);

DO:

前向传播:

x ← 残差单元1(*dataTrain*);

dataRes ← 残差单元 *N*(*x*);

x ← 带孔卷积1(*dataRes*);

dataDilated ← 带孔卷积 *N*(*x*);

x ← 全连接层1(*dataDilated*);

dataFC ← 全连接层 *N*(*dataDilated*);

反向传播:

使用Adam优化器执行梯度下降算法;

UNTIL 优化器目标的训练权重收敛或训练达到预定轮数;

IF 验证集精度 > 当前最佳精度:

保存当前模型权重;

END IF

END LOOP

6 实验与结果分析

本文融合残差神经网络和带孔卷积神经网络,实现了对八类交通模式的局部特征和全局特征提取,并通过多尺度特征提取,进行对交通模式的分类。本章通过对比实验验证了该算法的有效性。数据集区分过程,将使用的数据集,先分为五份,取其中一份作为测试集,剩下的四份作为训练集。

6.1 不同算法的对比实验

为了实现对比实验的公平性,对 *K* 最近邻(*K*-Nearest Neighbor, KNN)、LSTM、Alexnet、DenseNet201、Googlenet、SVM、随机森林、决策树和本文算法验证时,采用了相同的训练集和测试集进行实验。

在表3中记录了各类算法在八类交通模式的准确率、召回率和F1-分数三个评判标准。从实验结果上看,跑步和走路这两种模式是最容易区分的,这主要与人体的特殊运动方式有关,人们在走路和跑步时,都会有较大的震动和固定的运动频率,所以每一种算法在这两种模式上表现都较好。

从表3~5中可以看出,骑自行车这一交通模式在KNN和SVM上表现得不好,主要是由于这两种算法不能够很好地抓住骑自行车这一交通模式的有效特征;但是其他算法可以有效地抓住有效特征进行识别。从表4中可以看出,融合残差神经网络和带孔卷积神经网络(ResDilate)算法,在三种评价指标上表现最好,全部达到了95%以上,且准确率平均超过其他算法10个百分点。

从实验结果可以看出,公交车和私家车这两种交通模式,在各类算法中是比较容易混淆的。由于这两种交通模式,在特征上存在一定的相似性,尤其是在交通堵塞以及在高速路

上的情况下,公交车和私家车在速度、加速度以及拐弯等,这些特征都会非常相似,所以这两种交通模式识别存在一定的挑战性。Alexnet、DenseNet201 和 Googlenet 这三种算法,在识别这两种交通工具的三种评估下均可以达到 50% 以上。KNN、LSTM、随机森林、SVM 和决策树这五种算法在识别这两种交融模式的准确率在 30%~50%。从这个角度上看,深度学习算法的性能在一定程度上可以验证,深度学习提取的特征

比人工提取特征更具有识别性,更有利于识别交通模式。同理,火车和地铁在短时间内的特征也是非常相似的。使用单一的 Alexnet、DenseNet201 和 Googlenet 算法提取出的同一纬度的特征,不具有较强的识别性。使用人工提取出的特征,利用 SVM 等机器学习算法进行分类,也不能达到所需要的准确率。本文所提出的算法,在准确率、召回率和 F1-分数评估方法均可以达到 80% 以上。

表 3 不同算法的准确率比较

单位: %

Tab. 3 Accuracy comparison of different algorithms

unit: %

算法	静止	走路	跑步	自行车	私家车	公交车	火车	地铁
Alexnet	45.71	89.66	99.99	75.00	60.84	40.00	49.65	28.77
DenseNet201	55.03	90.48	99.99	55.96	53.71	48.65	59.50	27.27
Googlenet	52.55	90.41	99.99	70.19	61.11	41.50	57.55	24.73
LSTM	35.61	85.82	99.99	51.30	34.65	26.70	44.85	14.29
KNN	53.33	58.33	66.67	36.84	47.06	12.72	21.43	22.22
决策树	24.07	56.82	83.46	25.83	20.49	21.93	19.86	12.96
随机森林	33.33	75.81	91.89	41.48	32.85	29.41	35.53	23.62
SVM	39.41	42.04	53.09	41.33	39.22	33.33	36.55	14.02
ResDilate	88.61	95.42	99.99	97.90	90.11	94.70	81.90	91.43

表 4 不同算法的召回率比较

单位: %

Tab. 4 Recall comparison of different algorithms

unit: %

算法	静止	走路	跑步	自行车	私家车	公交车	火车	地铁
Alexnet	72.26	83.87	97.13	84.25	50.00	37.50	38.17	26.92
DenseNet201	67.10	85.81	94.83	83.56	70.69	37.50	38.71	19.23
Googlenet	66.45	85.16	98.28	77.4	50.57	42.36	43.01	29.49
LSTM	77.42	78.06	97.13	67.81	20.11	38.19	3.28	0.64
KNN	27.59	33.33	53.33	35.00	38.10	58.33	16.67	14.29
决策树	20.97	59.06	78.72	25.20	19.08	21.74	18.79	17.21
随机森林	42.74	74.02	96.45	59.35	34.35	21.74	18.12	24.59
SVM	36.49	43.71	48.31	39.49	43.17	40.27	36.55	12.10
ResDilate	90.32	94.19	98.85	95.89	94.25	86.81	92.47	82.05

表 5 不同算法的 F1-分数比较

单位: %

Tab. 5 F1-score comparison of different algorithms

unit: %

算法	静止	走路	跑步	自行车	私家车	公交车	火车	地铁
Alexnet	56.00	86.67	98.54	79.35	54.89	38.71	43.16	27.81
DenseNet201	60.47	88.08	97.35	67.03	61.04	42.35	46.91	22.56
Googlenet	58.69	87.71	99.13	73.62	55.35	41.92	49.23	26.90
LSTM	48.78	81.76	98.54	58.41	25.45	31.43	37.89	1.22
KNN	36.37	42.42	59.62	35.90	42.11	20.90	18.75	17.39
决策树	22.41	57.92	81.02	25.51	19.76	21.83	19.31	14.79
随机森林	37.46	74.90	94.12	48.83	33.58	25.00	24.00	24.10
SVM	37.89	42.86	50.59	40.39	41.10	36.47	36.55	12.99
ResDilate	89.46	94.81	99.42	96.89	92.13	90.58	86.87	86.49

6.2 本文算法的自对比实验

本文为了验证所提出算法各个部分的有效性,进行了自对比实验,实验结果如表 6~8 所示。

其中,表 6 是单一使用残差神经网络对交通模式进行分类。从残差神经网络的结构上可以看出,该网络结构可以对数据集中的局部特征进行有效提取,也就是说,可以将数据集中的各类交通模式上的短时间内的特征有效提取出来,因此在识别自行车、公交车、静止和火车上可以提取出有效的特征。从实验结果上看,这些短时间的局部特征比长时间的全局特征更加有效,准确率、召回率和 F1-分数均优于带孔卷积神经网络。

表 7 是使用带孔卷积神经网络实现交通模式识别实验结果,带孔卷积神经网络可以在较长时间内提取全局特征。从实验结果可以看出,对于火车这一交通模式,获得全局特征更有利于识别该交通模式。因为火车常常运行时间较长,而且加速度或者匀速的时间较长,在短时间内提取的局部特征,往往不具备很多的区分性,所以提取长时间的全局特征更加具有区分效果。

利用残差网络学习特征图像,可以将感受野控制在相对较小的范围内,从而学习局部更细粒度的特征表达;同理,利用带孔卷积的卷积核间隔,在不扩大参数规模的前提下,增加感受野,忽略掉一定的局部特征但是可以学习到全局特征。

表 6 使用残差网络的实验结果 单位: %
Tab. 6 Experimental results of utilization of residual network unit: %

交通模式	准确率	召回率	F1-分数
静止	49.21	60.65	54.34
走路	94.29	85.16	89.49
跑步	99.99	95.98	97.95
自行车	64.77	78.08	70.81
私家车	61.11	44.25	51.33
公交车	32.57	39.58	35.74
火车	46.49	46.24	46.36
地铁	24.62	20.51	22.83

表 7 使用带孔卷积神经网络的实验结果 单位: %
Tab. 7 Experimental results of utilization of dilated convolutional network unit: %

交通模式	准确率	召回率	F1-分数
静止	50.20	80.00	61.69
走路	91.95	88.39	90.13
跑步	99.99	97.70	98.84
自行车	74.14	88.36	80.63
私家车	68.90	64.94	66.86
公交车	68.69	47.22	55.97
火车	51.70	48.92	50.28
地铁	36.04	25.64	29.96

综合上面的实验结果可知,本文算法将两种神经网络进行融合,就可以得到表 8 中的识别效果。如表 8 所示,残差网

络和带孔卷积融合后,八类交通模式识别精度均有一定的提高,可以验证本文设计算法的有效性和两种算法融合的必要性。

表 8 融合残差网络和带孔卷积神经网络的实验结果 单位: %
Tab. 8 Experimental results of fusion of residual network and dilated convolutional network unit: %

交通模式	准确率	召回率	F1-分数
静止	88.61	90.32	89.46
走路	95.42	94.19	94.81
跑步	99.99	98.85	99.42
自行车	97.90	95.89	96.89
私家车	90.11	94.25	92.13
公交车	94.70	86.81	90.58
火车	81.90	92.47	86.87
地铁	91.43	82.05	86.49

6.3 调参实验

在本文中,主要调整了模型中的超参数,其中包括学习率、正则化参数值,以及残差神经网络卷积核大小。其中本文使用效果最好的参数已经在表 1 和表 2 给出。

本文调整的其他参数如表 9 所示。通过控制变量法的原则,每次调整一个参数。从各个交通模式的准确率可以看出,超级参数的调整,对于准确率影响非常大,尤其是对于公交车和私家车、地铁和火车,这些比较难以区分的交通模式有明显准确率变化。

表 9 不同参数的准确率比较 单位: %
Tab. 9 Accuracy comparison of different parameters unit: %

参数	交通模式							
	静止	走路	跑步	自行车	私家车	公交车	火车	地铁
学习率 0.50, 卷积核(5,5,3), 正则化值(0.0001)	52.97	94.07	99.99	71.18	58.67	69.51	42.31	43.42
学习率 0.50, 卷积核(3,3,2), 正则化值(0.0001)	58.10	91.33	99.99	79.87	67.05	55.91	51.28	34.51
学习率 0.50, 卷积核(3,3,2), 正则化值(0.0005)	56.18	92.00	99.99	82.98	60.40	62.86	50.31	40.11
学习率 0.25, 卷积核(3,3,2), 正则化值(0.0005)	51.65	90.67	99.99	78.34	67.84	63.55	53.93	35.65
学习率 0.25, 卷积核(3,3,2), 正则化值(0.0001)	52.61	93.15	99.99	81.46	66.87	64.60	50.77	38.66
学习率 0.25, 卷积核(5,5,3), 正则化值(0.0001)	53.78	95.71	98.84	72.12	60.38	62.86	52.27	35.79
学习率 0.25, 卷积核(5,5,3), 正则化值(0.0005)	53.23	94.41	99.99	79.22	64.29	49.65	54.39	33.33
学习率 0.50, 卷积核(5,5,3), 正则化值(0.0005)	88.61	95.42	99.99	97.90	90.11	94.70	81.90	91.43

7 结语

本文通过融合残差网络和带孔卷积,对八类交通模式识别进行识别。首先,通过 FFT 算法将一维传感器时域数据转换二维频域图像数据;然后,使用 PCA 算法有效降低图像数据复杂度;接着,使用残差网络提取交通模式局部特征,使用带孔卷积网络提取全局特征;最后,通过对局部特征和全局特征的深度融合,实现了对交通模式的准确识别,验证了融合残差网络和带孔卷积网络模型的准确性和有效性。后续工作中将着力于继续提升模型精度并进一步降低模型复杂度,从而更好地满足低功耗应用需求。

参考文献 (References)

[1] WANG C, LUO H, ZHAO F, et al. Combining residual and LSTM recurrent networks for transportation mode detection using multimodal sensors integrated in smartphones [J]. IEEE Transactions on Intelligent Transportation Systems, 2020 (Early Access): 1-13.

[2] QIN Y, LUO H, ZHAO F, et al. Toward transportation mode recognition using deep convolutional and long short-term memory recurrent neural networks [J]. IEEE Access, 2019, 7: 142353-142367.

[3] WIDHALM P, NITSCHKE P, BRÄNDIE N. Transport mode detection with realistic smartphone sensor data [C]// Proceedings of the 21st International Conference on Pattern Recognition. Piscataway: IEEE, 2012: 573-576.

[4] ZHENG Y, LIU L, WANG L, et al. Learning transportation mode from raw GPS data for geographic application on the web [C]// Proceedings of the 2008 17th International Conference on World Wide Web. New York: ACM, 2008: 247-256.

[5] ENDO Y, TODA H, NISHIDA K, et al. Deep feature extraction from trajectories for transportation mode estimation [C]// Proceedings of the 2016 Pacific-Asia Conference on Knowledge Discovery and Data Mining, LNCS 9652. Cham: Springer, 2016: 54-66.

- [6] JIANG X, DE SOUZA E N, PESARANGHADER A, et al. TrajectoryNet: an embedded GPS trajectory representation for point-based classification using recurrent neural networks [C]// Proceedings of the 2017 27th Annual International Conference on Computer Science and Software Engineering. Armonk: IBM, 2017: 192-200.
- [7] GRAVES A, JAITLY N. Towards end-to-end speech recognition with recurrent neural networks [C]// Proceedings of the 31st International Conference on Machine Learning Research. New York: JMLR.org, 2014: 1764-1772.
- [8] 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 Symposium on Advances in Geographic Information Systems. New York: ACM, 2011: 54-63.
- [9] BEDOGNI L, DI FELICE M, BONONI L. By train or by car? Detecting the user's motion type through smartphone sensors data [C]// Proceedings of the 2012 IFIP Wireless Days. Piscataway: IEEE, 2012: 1-6.
- [10] JAHANGIRI A, RAKHA H A. Applying machine learning techniques to transportation mode recognition using mobile phone sensor data [J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(5): 2406-2417.
- [11] FANG S H, FEI Y X, XU Z, et al. Learning transportation modes from smartphone sensors based on deep neural network [J]. IEEE Sensors Journal, 2017, 17(18): 6111-6118.
- [12] 王璞, 刘洋, 黄智仁. 一种轻量级梯度提升机的交通模式识别 [J]. 哈尔滨工业大学学报, 2019, 51(9): 96-102. (WANG P, LIU Y, HUANG Z R. Transportation modes recognition using a light gradient boosting machine [J]. Journal of Harbin Institute of Technology, 2019, 51(9): 96-102.)
- [13] 王昊, 刘高军, 段建勇, 等. 基于特征自学习的交通模式识别研究 [J]. 哈尔滨工程大学学报, 2019, 40(2): 354-358. (WANG H, LIU G J, DUAN J Y, et al. Transportation mode detection based on self-learning of features [J]. Journal of Harbin Engineering University, 2019, 40(2): 354-358.)
- [14] 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.)
- [15] YANG J B, NGUYEN M N, SAN P P, et al. Deep convolutional neural networks on multichannel time series for human activity recognition [C]// Proceedings of the 2015 24th International Joint Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2015: 3995-4001.
- [16] ZHOU P, SHI W, TIAN J, et al. Attention-based bidirectional long short-term memory networks for relation classification [C]// Proceedings of the 2016 54th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2016: 207-212.
- [17] XIAO Z, WANG Y, FU K, et al. Identifying different transportation modes from trajectory data using tree-based ensemble classifiers [J]. ISPRS International Journal of Geo-Information, 2017, 6(2): Article No. 57.
- [18] FENG T, TIMMERMAN H J P. Comparison of advanced imputation algorithms for detection of transportation mode and activity episode using GPS data [J]. Transportation Planning and Technology, 2016, 39(2): 180-194.
- [19] IGNATOV A. Real-time human activity recognition from accelerometer data using convolutional neural networks [J]. Applied Soft Computing, 2018, 62: 915-922.
- [20] LIU H, LEE I. End-to-end trajectory transportation mode classification using Bi-LSTM recurrent neural network [C]// Proceedings of the 2017 12th International Conference on Intelligent Systems and Knowledge Engineering. Piscataway: IEEE, 2017: 1-5.
- [21] HANNUN A, CASE C, CASPER J, et al. DeepSpeech: scaling up end-to-end speech recognition [EB/OL]. [2020-04-26]. <https://arxiv.org/pdf/1412.5567v1.pdf>.
- [22] BOHTE W, MAAT K. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: a large-scale application in the Netherlands [J]. Transportation Research Part C: Emerging Technologies, 2009, 17(3): 285-297.
- [23] FRIEDRICH B, LÜBBE C, HEIN A. Combining LSTM and CNN for mode of transportation classification from smartphone sensors [C]// Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers. New York: ACM, 2020: 305-310.

This work is partially supported by the National Natural Science Foundation of China (61872046), the Scientific and Technological Innovation Ability Improvement Action Program of Beijing University of Posts and Telecommunications (2019XD-A06).

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

QIN Yanjun, born in 1991, Ph. D. candidate. Her research interests include urban computing, transportation mode recognition.

WANG Chenxing, born in 1997, M. S. candidate. His research interests include urban computing, transportation mode recognition, traffic flow forecasting.

GAO Cunyuan, born in 2000. His research interests include urban computing, transportation mode recognition.

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.

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