

引文格式:李静,刘海砚,郭文月,等.基于深度学习的人群活动流量时空预测模型[J].测绘学报,2021,50(4):522-531. DOI:10.11947/j. AGCS.2021.20200230.
LI Jing, LIU Haiyan, GUO Wenyue, et al. A spatio-temporal network for human activity prediction based on deep learning[J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(4): 522-531. DOI: 10.11947/j. AGCS.2021.20200230.

基于深度学习的人群活动流量时空预测模型

李 静¹, 刘海砚¹, 郭文月², 陈 欣²

1. 信息工程大学数据与目标工程学院, 河南 郑州 450052; 2. 信息工程大学地理空间信息学院, 河南 郑州 450052

A spatio-temporal network for human activity prediction based on deep learning

LI Jing¹, LIU Haiyan¹, GUO Wenyue², CHEN Xin²

1. Institute of Data and Target Engineering, Information Engineering University, Zhengzhou 450052, China;
2. Institute of Surveying and Mapping, Information Engineering University, Zhengzhou 450052, China

Abstract: The traditional spatio-temporal prediction methods could hardly model the complex nonlinear relationship of spatio-temporal phenomena, thus they lack the ability to consider the influence of spatial multi-scale characteristics into the prediction results. In order to overcome this deficiency, a novel model of space-time network (MST-Net) is proposed in this paper, which transforms the regression problem of volume prediction into a discriminant model with time-space characteristics. The spatial and temporal characteristics of spatio-temporal data are extracted by multi-scale parallel convolution and gate recurrent unit respectively. Thus the extracted features are fused with the attention mechanism introduced to capture the long-term features. Finally, the prediction results can be obtained by using the full connection layers. In order to prove the reliability and validity of the model, the model is tested on two challenging social media sign-in datasets. The results indicate that the proposed model outperformed other algorithms in two prediction results evaluation indexes, namely the root mean square errors (RMSE) and mean absolute percentage errors (MAPE), which illustrate that the proposed method could achieve higher prediction accuracy and could better fit the nonlinear relationship of the space-time problem. The proposed model is suitable to predict the flow of human activities.

Key words: spatial multiscale; spatio-temporal network; spatio-temporal prediction; parallel convolution; gate recurrent unit

Foundation support: The National Natural Science Foundation of China (No. 41801388); Natural Science Foundation of Henan Province (No. 182300410005)

摘 要:传统的时空预测方法缺乏对复杂时空非线性关系的描述,且难以顾及空间多尺度特征对于预测结果的影响。针对这一问题,本文提出了一种融合空间多尺度特征的时空网络模型(MST-Net),将流量预测的回归问题转换为具有时空特性的判别模型。首先,通过并联卷积提取空间多尺度特征;然后,通过引入注意力机制的门控循环单元提取时间特征;最后,利用全连接层得到预测结果。本文将该模型用于人群活动流量的预测,分别在两组真实的社交媒体签到数据集上进行试验。试验结果表明:本文采用的卷积层连接方式和特征融合方法,相比于单层卷积层提取空间特征、其他连接方式和融合方法以及传统的时空预测模型,在均方根误差(RMSE)和平均绝对百分比误差(MAPE)两个预测结果评价指标上均有不同程度的提高,说明本文方法具有较高的预测精度,能够较好地拟合时空问题的非线性关系,实现人群活动流量的预测。

关键词:空间多尺度;时空网络;时空预测;并联卷积;门控循环单元

中图分类号:P208

文献标识码:A

文章编号:1001-1595(2021)04-0522-10

基金项目:国家自然科学基金(41801388);河南省自然科学基金(182300410005)

人群活动流量预测是探索人类移动规律的重要环节,也是时空预测的核心问题之一^[1]。对人群活动流量的预测研究能够用来分析城市的交通状况,帮助理解城市的功能区域及人口活动分布等,在交通运维^[2]、灾难响应^[3]、旅游推荐^[4]、城市规划^[5]等方面均具有重要应用价值。当前对于该问题的研究主要是在对地理区域进行基本单元划分的基础上,通过建立区域间人群流量的时空关联来实现对人群活动流量的预测。然而,这种时空关联包含了人群活动的时间依赖性和空间依赖性,是相对复杂的非线性数学模型。因此,如何对这种复杂的时空关联进行构建是该研究领域的主要挑战之一。

早期研究侧重关注时间依赖性,即区域内人群活动流量随时间的动态变化,包括周期性和趋势性,常用的方法有时间序列模型^[6]、线性回归模型^[7]和卡尔曼滤波模型^[8]等。文献[9]使用历史平均模型(historical average, HA),将历史时期交通量的平均值作为预测值,该方法不需要任何假设,计算简单快速,但没有顾及时间特征,且预测精度相对较低。文献[10]使用非参数化的自回归移动平均模型(autoregressive integrated moving average model, ARIMA)对兴趣点的人流量进行时序建模,但不能很好地拟合非平稳性历史时序数据。以上方法均只考虑了时间依赖性,忽略了空间特征对于人群活动规律的影响。空间依赖性是指人群活动与空间之间的相互影响,例如上游道路交通状态通过传递效应影响下游道路交通状态,下游道路交通状态通过反馈效应影响上游道路交通状态^[11]。文献[12]根据区域人流情况的相似性,将相邻的城市格网区域聚集来得到城市的功能区域,进而用核密度估计的方式预测每个区域的人流量分布。文献[13]对道路网使用矩阵分解来学习道路连接区域之间的潜在空间以预测流量。这些方法在时间依赖的基础上考虑了空间依赖关系,但却没有将二者同时结合起来,因此无法很好地拟合复杂的时空非线性关系。

由于深度神经网络对于不连续、非线性的问题具有更好的性能^[14],近些年也逐渐被学者用于时空预测问题上,通过循环神经网络(recurrent neural network, RNN)提取时间维度的特征,通

过卷积神经网络(convolutional neural network, CNN)提取空间维度特征,然后将二者结合构建时空网络模型。文献[15]同时构造3个相同的卷积结构来提取人群活动时间特征的趋势性、周期性和邻近性。文献[16]提出了一种用于需求预测的多视点时空网络,该网络通过集成长短记忆网络(long short-term memory, LSTM)、局部卷积神经网络和语义网络嵌入,同时学习时空相关性。文献[17]采用图神经网络用于学习复杂的拓扑结构来捕获空间依赖关系,同时考虑了时间尺度对预测结果的影响。文献[18]将LSTM与注意力机制(attention mechanism)结合的方法用来同时提取时间的短期和长期特征。以上方法尽管能够拟合时空非线性关系,但这些方法多是基于静态的、单一尺度的空间特征,难以顾及动态的空间交互特征以及空间多尺度特征对于深度学习模型的学习能力和预测结果的影响。

针对上述问题,本文提出一种融合空间多尺度特征的时空网络预测模型(multi-scale characteristics spatio-temporal network, MST-Net),将静态的人群流量和动态的人群交互流作为卷积神经网络输入来提取空间特征,同时使用并联卷积进行空间多尺度特征融合,随后使用门控循环单元(gate recurrent unit, GRU)来提取时间特征。试验结果证明,本文提出的时空网络模型能够提高预测精度和学习效率,实现人群活动流量预测,为感知人类的时空移动规律提供方法支持。

1 融合空间多尺度特征的时空预测模型

时空预测不仅包含时间序列的分析,同样也受到实体空间特征的影响。基于这一假设,本文提出的MST-Net模型主体由两部分组成:①通过局部卷积神经网络提取不同尺度下的空间特征来描述空间依赖性;②通过门控循环单元提取时间特征来描述时间依赖性,网络主要结构如图1所示。首先,为使网络更好地学习位置间的空间依赖性,网络中同时顾及区域内人群活动静态流量和区域间交互人群流量,再将局部卷积神经网络提取的不同尺度下的空间特征进行融合。然后,将融合后的特征输入门控循环单元中。时间序列数据不仅具有短期依赖性还具有一定的周期

性,即长期依赖性,通过 GRU 中单元间的信息传输能够实现长时与短时间特征的提取与表征。此外,网络对于所提取的特征是按照等权的处理方式参与运算,具有显著性的特征则难以在网络中发挥更大作用,基于此,本文进一步引入注意力机

制^[19]来提取流量数据的长短周期性以增强 GRU 所提取特征的显著性,从而提升时空预测任务的精度。最后,将学习到的时间特征向量输入到全连接层,经过激活函数进行回归运算,从而得到下一个时刻的流量值。

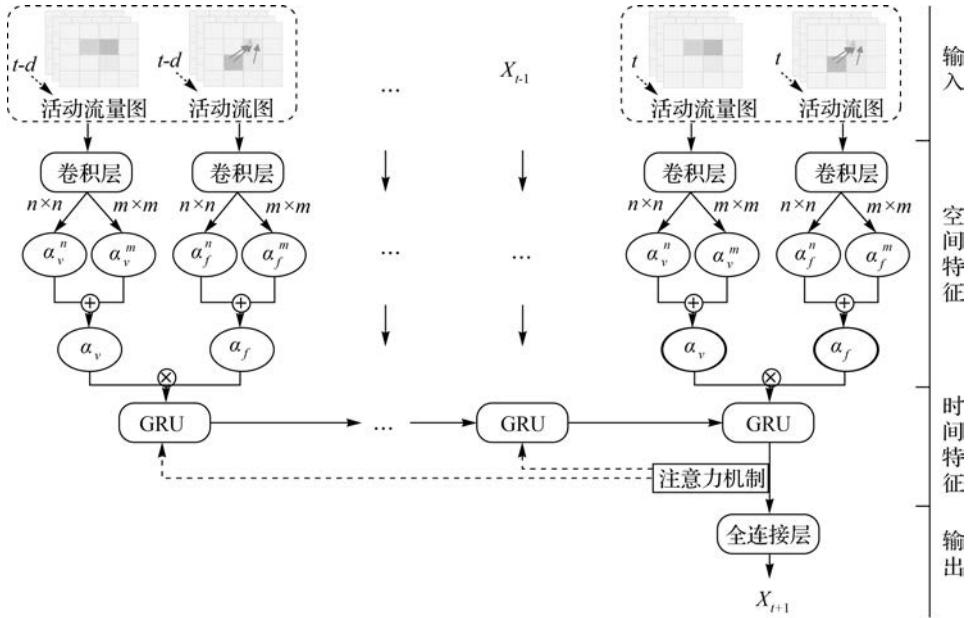


图 1 MST-Net 模型结构
Fig.1 The network structure of MST-Net

1.1 基于局部卷积神经网络的静态和动态空间特征提取

根据地理学第一定律可知,人群活动在空间上存在一定的相关性^[20],区域的人群活动受其邻域空间变量影响,而卷积神经网络能够通过卷积运算捕获局部的空间特征^[21];对于时空预测问题,下一时刻的流量依赖该区域的历史流量,区域间的人群流动能够强化区域间的动态空间关系^[22]。基于此,本文采用卷积神经网络提取空间特征,通过规则格网划分将整个研究区域转换为规则格网,将格网内流量视为对应像素的灰度,从而将其转换为图像,利用时间间隔划分人群流量图和人群交互流图作为网络输入,通过人群流量图提取静态空间特征,通过人群交互流图提取动态空间特征。

为了方便空间关系表达,本文采用文献[18]的空间划分方式将区域进行规则格网划分,研究区域 G 在时间段 T 内的时空关系。假设区域被划分为 n 个格网单元 $G = \{g_1, g_2, \dots, g_i, \dots, g_n\}$,将时间段 T 内 m 个等步长时间间隔 $T = \{t_1, t_2, \dots, t_i, \dots, t_m\}$ 内,定义如下时空参数以方

便后续描述。

定义 1:区域内人群活动流量 $V_{g_i}^t$ 。定义在一个时间间隔 t 内,人群活动流量为经过该格网区域 g 的人群数量。

定义 2:区域间人群活动流 $F_{(g_i, g_j)}^t$ 。定义在一个时间间隔 t 内,从格网单元 g_i 转移到格网单元 g_j 的人群数量,其中, g_i 为流的起始区域, g_j 为流的终止区域,如图 2(a)所示。

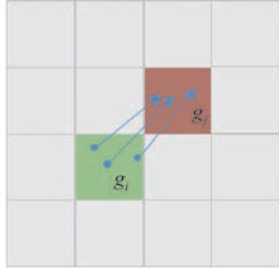
定义 3:人群活动输入流 $F_{in(g_i)}^t$ 。定义人群活动输入流为在时间间隔 t 内,其他格网单元流入格网单元 g_i 的人群数量,如图 2(b)所示。

定义 4:人群活动输出流 $F_{out(g_i)}^t$ 。定义人群活动输出流为格网单元 g_i 在时间间隔 t 内,流向其他格网单元的人群数量,如图 2(c)所示。

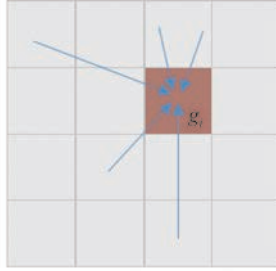
为提取目标区域 g_i 的空间特征,对于一个时间间隔 t ,使用局部人群流量图 $P_{g_i}^v$ 和局部人群活动流图 $P_{g_i}^f$ 作为网络输入。假设目标区域 g_i 的局部图像尺寸为 $r \times r \times 1$,且目标区域位于局部图像的中心,则 $P_{g_i}^v$ 的图像尺寸为 $r \times r \times 1$, $P_{g_i}^f$ 的图像尺寸为 $r \times r \times 2$,其中 2 代表 $P_{g_i}^f$ 的两个通道,一

个通道为局部人群输入流图,一个通道为局部人群输出流图。因此,局部卷积神经网络的输出为

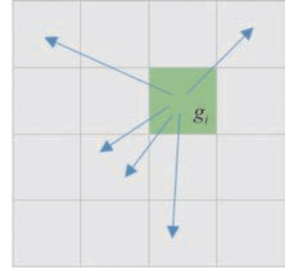
$$\left. \begin{aligned} \alpha_v &= F(W_v^t P_v^t + b_v^t) \\ \alpha_f &= F(W_f^t P_f^t + b_f^t) \\ \alpha_r &= \alpha_v \otimes \alpha_f \end{aligned} \right\} \quad (1)$$



(a) 从格网区域 g_i 到格网区域 g_j 人群活动流



(b) 格网单元 g_i 的输入流



(c) 格网单元 g_i 的输出流

图 2 基于格网单元的交互流、输入流与输出流

Fig.2 Interaction flow, input flow and output flow based on grid cell

1.2 基于并联卷积的空间多尺度特征融合

尺度是地理信息科学最重要的话题之一^[23]。不同空间尺度对模型的预测精度有一定影响。在卷积网络中,每一层输出的特征图上的像素点在输入图片上映射的区域大小称为感受野(receptive-field)^[24]。在图像大小一定的情况下,卷积核的尺寸决定了感受野的大小^[25],从而决定了图像内提取特征的范围,对应于参与运算的地理范围,即卷积核对应的空间尺度。由图 3 可知,当图像比例尺确定时,卷积核尺寸越大对应的视觉感受野越大,特征图能表示的原始图像范围越大,越能表现原始图像更大尺度的空间特征;当卷积核越小,其对应的感受野越小,特征图表示的原始图像范围越小,越能表现图像的局部特征^[26]。此外,随着卷积层数的增加,其感受野也越大。例如,图 3 中尺寸为 5×5 的卷积核与尺寸为 3×3 的卷积核串联,其对应的感受野等效于一个尺寸为 7×7 卷积核的感受野。

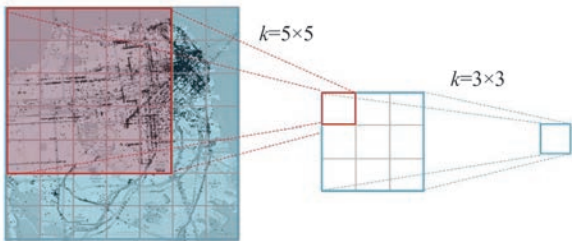


图 3 当卷积核步长为 1 时不同卷积核对应的感受野

Fig.3 The receptive field corresponding to different convolution kernels when the stride parameter equals 1

式中,符号“ \otimes ”表示张量乘积; α_v 表示静态空间特征; α_f 表示动态空间特征; α_r 在局部图像尺度为 $r \times r$ 上的空间特征;采用 $F = \max(0, x)$ 线性整流函数 ReLU 作为激活函数; W_v^t, W_f^t 为卷积层的权重向量; b_v^t, b_f^t 为卷积层的偏置项。

本文通过控制卷积核的大小来实现不同空间尺度特征的提取,同时使用并联卷积的方式将不同大小卷积核提取的多维特征向量扁平化为一维特征向量后进行拼接输入下一层网络中,实现不同空间尺度下的特征融合,如图 4 所示。

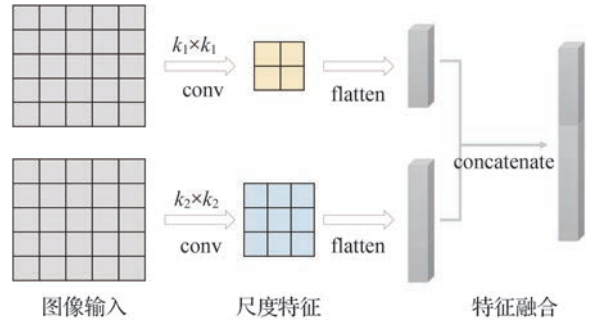


图 4 基于并联卷积的空间多尺度特征融合

Fig.4 Spatial multiscale feature fusion based on parallel convolution

经过并联卷积融合后的输出特征可表示为

$$\left. \begin{aligned} \alpha_{v(k_1, k_2, \dots, k_n)} &= F(W_v^n [P_{v, k_1}^t | P_{v, k_2}^t | \dots | P_{v, k_n}^t] + b_v^t) \\ \alpha_{f(k_1, k_2, \dots, k_n)} &= F(W_f^n [P_{f, k_1}^t | P_{f, k_2}^t | \dots | P_{f, k_n}^t] + b_f^t) \\ \alpha_{r(k_1, k_2, \dots, k_n)} &= \alpha_{v(k_1, k_2, \dots, k_n)} \otimes \alpha_{f(k_1, k_2, \dots, k_n)} \end{aligned} \right\} \quad (2)$$

式中,符号“ $|$ ”表示按行拼接; k 表示不同大小的卷积核; $\alpha_{v(k_1, k_2, \dots, k_n)}$ 表示不同尺度下融合后的静态空间特征; $\alpha_{f(k_1, k_2, \dots, k_n)}$ 表示不同尺度下融合后的动态空间特征; $\alpha_{r(k_1, k_2, \dots, k_n)}$ 在局部图像尺度为 $r \times r$ 的不同尺度融合后的空间特征;采用 $F = \max(0, x)$ 线性整流函数 ReLU 作为激活函数; W_v^n, W_f^n 为卷积层的权重向量; b_v^t, b_f^t 为卷积层的偏置项。

1.3 基于门控循环单元的时间特征提取

目前,应用最广泛的处理序列数据的神经网络是递归神经网络(RNN)^[27]。但是,传统的RNN在处理长期依赖问题上存在梯度消失和梯度爆炸等问题^[28]。长短时记忆网络(LSTM)^[29]和门控循环单元(GRU)^[30]作为RNN的改进,通过增加门控机制被证实已能够很好解决长期依赖的问题。LSTM和GRU在训练效果上都有很好的表现,但是由于LSTM结构相对复杂,需要的训练时间较长,且训练参数较多。因此,本文选用结构相对简单,需要训练时间较短且参数较少的GRU来进行时间特征的提取。如图5所示, r' 是重置门,用于忽略前一时刻信息的程度; u' 是更新门,用于控制前一时刻信息进入当前时刻的程度; c' 为当前时刻的候选隐藏状态;GRU将 $t-1$ 时刻的隐藏状态 h^{t-1} , t 时刻的空间特征 α' 作为 t 时刻的输入,从而获取当前时刻的信息。

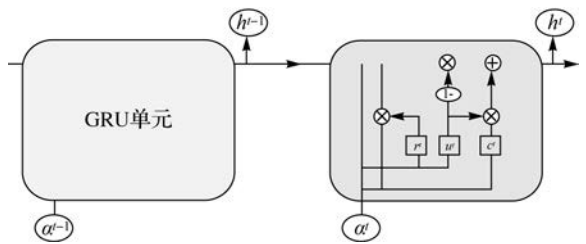


图5 门控循环单元模型结构
Fig.5 The structure of GRU

2 试验及分析

2.1 试验数据集与数据预处理

本文选取纽约市曼哈顿区人群活动签到数据集(NY)和旧金山市人群活动签到数据集(SFO)两个真实数据集对模型的预测性能进行测试,数据来源于Foursquare^[31]社交媒体签到平台。原始数据包含如图6所示的签到时间、签到经纬度等7个字段。根据本文模型,首先需要将矢量的签到点数据通过格网划分、时间间隔划分成时间序列图像后再输入模型,具体流程如图6所示。对不同时间间隔的人群流量进行统计,能够得到人群流量时间序列图像;对于相邻时间间隔用户的区域转移情况进行统计,能够得到人群输入输出流的时间序列图像。具体的人群输入输出流的计算原理如图7所示。假设用户 I 在 $t-1$ 时段处于 g_i ,在 t 时段处于 g_j ,则说明用户 I 在 t 时段由区域 g_i 转移至 g_j 。对于区域 g_j 在 t 时段有一个用户进入,区域 g_i 在 t 时段有一个用户流出。通过用户编码可以锁定一个用户在不同时段所处区域,从而可以统计得到某一区域在某一时段用户输入输出情况,进而量化为人群输入输出流的时间序列图像。

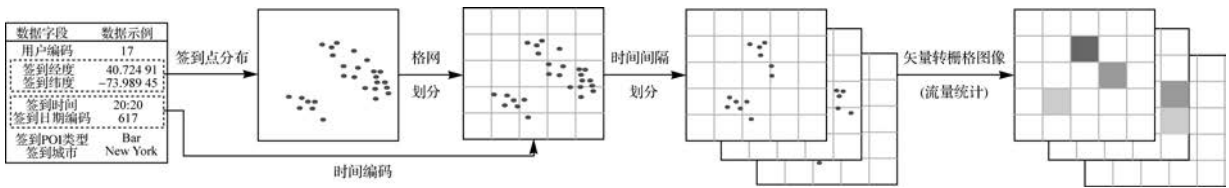


图6 原始数据详情及数据预处理流程
Fig.6 Details of raw data and data preprocessing process

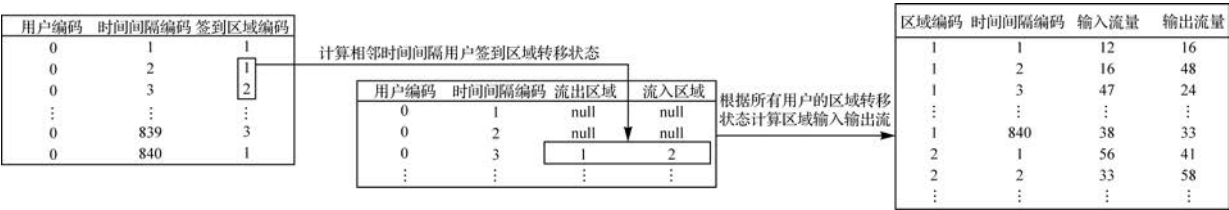


图7 人群活动输入输出流计算过程
Fig.7 Calculation process of crowd activity input and output flow

试验分别选取两个数据集2012年1月1日至2012年10月7日共280d的签到数据,选取前168d的数据作为训练集、中间42d的数据作为验证集,后70d作为测试集。顾及人群签到数据

采样时间较长且稀疏,若时间间隔划分过小,则无法展现数据的显著性;若时间间隔过大,则无法反映数据的周期性。因此,为了避免数据过于稀疏,且考虑到时间语义信息,本文将一天的 24 h 划分为 4 个时间间隔,即凌晨(0:00—06:00)、上午

(06:00—12:00)、下午(12:00—18:00)和晚上(18:00—24:00)。参照文献[18]的格网划分方式,研究将 NY 数据集划分为 20 行×10 列的规则格网,将 SFO 数据集划分为 20 行×20 列的规则格网。具体试验数据量及数据安排见表 1。

表 1 试验数据量及训练、验证、测试数据安排

Tab.1 The amount of experimental data and the arrangement of training and testing data						
数据集	天数/d	时间间隔 /6 h	格网 大小	训练集 数据量	验证集 数据量	测试集 数据量
NY	280	1120	20 行×10 列	29 885	7471	19 914
SFO	280	1120	20 行×20 列	29 108	7277	9602

2.2 模型训练及参数设定

试验的批大小设置为试验设备的显存上限 256,加速训练过程。学习率采用常用值 0.001^[17-18]。GRU 隐藏单元数会影响模型的性能和预测精度,研究分别将单元数设置为 16、32、64、100 和 128,并利用式(3)计算不同单元数对应的预测误差,结果如图 8 所示,当单元数为 128 时对应的预测误差最小,故将隐藏单元数设置为 128。试验将迭代次数设置为 55,由图 9 可知,模型收敛正常。

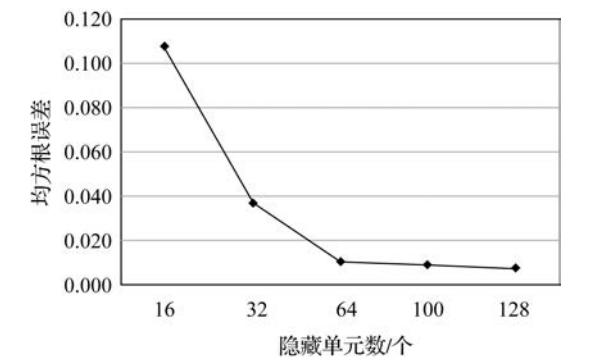


图 8 不同隐藏单元数对应的预测误差
Fig.8 Prediction errors under different hidden units

模型采用局部卷积神经网络提取空间特征并将每一个邻域大小设置为 7×7。卷积核的尺寸、卷积层数量和连接方式都会影响学习能力和预测结果^[32]。为了验证本文所提出的融合空间多尺度特征的时空网络模型的有效性,试验设置了 3×3 与 5×5 两种卷积核尺寸,并设计了 4 种不同的空间特征提取方案:①一层 64 个 5×5 卷积核的卷积层直接提取特征;②两层卷积层串联,即先经过 32 个 3×3 卷积核,再经过 32 个 5×5 卷积核提取特征;③两层卷积层并联后经过乘积运算融

合特征,即将经过 32 个 3×3 卷积核直接提取的特征与经过 32 个 5×5 卷积核直接提取的特征通过乘积运算得到融合后的特征;④两层卷积层并联后经过拼接融合特征,即将经过 32 个 3×3 卷积核直接提取的特征与经过 32 个 5×5 卷积核直接提取的特征通过拼接得到融合后的特征。其中,方案④为本模型使用的融合方案。

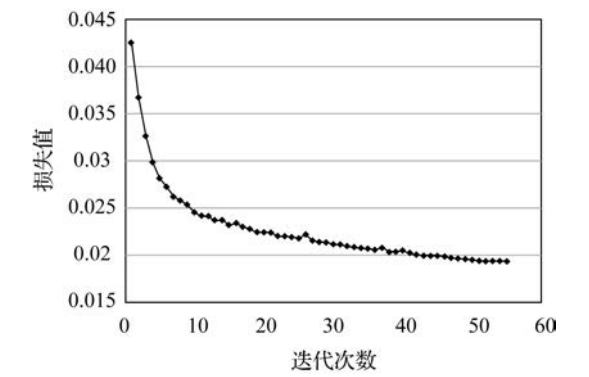


图 9 损失值随迭代次数增加的变化曲线
Fig.9 Evolution of loss value with increasing number of epochs

2.3 试验结果及分析

2.3.1 模型精度评价指标

试验选用常用的模型预测结果衡量指标均方根误差(RMSE)和平均百分比误差(MAPE)来对模型精度进行评价。均方根误差能够衡量观测值与真实值之间的偏差,平均百分比误差能够考虑预测值与真实值的误差,以及误差与真实值之间的比例。二者的值越小说明模型的精度越高。其计算公式如下

$$RMSE = \sqrt{\frac{1}{\xi} \sum (\hat{y}_{i+1}^i - y_{i+1}^i)^2}$$

(3)

$$MAPE = \frac{1}{\xi} \sum_{i=1}^{\xi} \frac{|\hat{y}_{t+1}^i - y_{t+1}^i|}{y_{t+1}^i} \quad (4)$$

式中, y_{t+1}^i 与 \hat{y}_{t+1}^i 分别表示下一时刻 $t+1$ 区域 i 人群活动流量的真实值与预测值; ξ 表示样本总数。

2.3.2 不同空间尺度和融合方式的精度对比

根据 2.2 节中设计的 4 种空间特征提取方案分别在两个数据集上进行试验,结果见表 2。通过对比 4 种方法在两个数据集上的试验结果发现,本文方法在两个评价指标上都具有更高的精度,主要结论为:

(1) 从卷积核数量来看,单一卷积核精度较多卷积核精度低,说明网络通过增加卷积核及卷积层数量能够提高模型的准确率和稳定性。

(2) 从卷积层连接方式来看,本文中采用卷积层并联的方式比串联的方式百分比误差小 1.9%~6.4%,这是因为串联方式增大了感受野(原理如图 3 所示),虽然强化了全局特征,但缺失了局部特征。而并联方式同时提取了不同尺度下的空间特征,包含全局特征和局部特征,提升特征的丰富度,从而强化了下一层网络的输入,使输出精度提高。

(3) 从并联卷积层的特征融合方式对比,采用特征拼接比特征乘积方式百分比误差小 3.9%~4.4%,特征乘积后会导致特征间的线性关系改变,从而无法完全保留原有特征的特性,而使用拼接方式既能够对特征进行融合,同时也保持了特征间的线性关系,是一种更为合理的融合方式。

此外,通过对比,在相同方案下 NY 数据集比 SFO 数据集的预测误差要小 19.9%~20.9%。说明在模型结构相同的情况下,数据的特性对于预测的结果也有较大的影响。在数据获取条件相同的情况下,主要考虑数据稀疏度和密度。研究将数据的稀疏度定义为无人群流量的空白格网单元数占总格网数的比例,利用式(5)对数据的稀疏度进行计算

$$s = \frac{n_{v=0}}{n} \times 100\% \quad (5)$$

利用式(6)计算除去空白格网后单元格网的数据密度

$$\rho = \frac{\sum v}{n - n_{v=0}} \quad (6)$$

式中, $n_{v=0}$ 表示人群流量为 0 的空白格网数; n 代

表研究区域的格网总数; v 代表单元格网区域的人群流量。

计算结果见表 3。从表 3 可以看出,NY 数据集比 SFO 数据集更稀疏,且单元格网的数据密度更大,这说明 NY 数据集存在更多无数据的空白区域,使得数据在空间分布上更加聚集,这种聚集效应增加了空间特征的显著性,使得数据的特征明显。

表 2 不同空间特征提取方案的精度比较

Tab.2 Precision comparison of different spatial feature extraction schemes

数据集	卷积核尺寸	连接方式	均方根误差	平均百分比误差/(%)
NY	(5×5)	—	0.007 3	15.28
	(3×3),(5×5)	串联	0.006 0	9.27
	(3×3),(5×5)	并联(乘积)	0.005 8	7.31
	(3×3),(5×5)	并联(拼接)	0.002 6	2.92
SFO	(5×5)	—	0.001 4	35.25
	(3×3),(5×5)	串联	0.001 1	30.16
	(3×3),(5×5)	并联(乘积)	0.001 0	27.72
	(3×3),(5×5)	并联(拼接)	0.000 9	23.87

表 3 数据稀疏度及密度

Tab.3 Data sparsity and density

数据集	空白格网单元数	格网总数	数据稀疏度/(%)	单元格网数据密度
NY	134	200	67.00	867.71
SFO	97	400	24.25	199.93

研究选取某一格网单元在特征提取方案①和方案④(本文模型所采用的融合方案)的真实值与预测值进行了可视化,如图 10 所示。发现在相同数据集上方案④的拟合效果更好,在不同数据集上,NY 数据集的拟合效果更好。对于 SFO 数据集,在流量峰值附近的预测值与真实值误差较大,说明模型对于峰值的预测有所欠缺。同时,对选取两个数据集上一天中 4 个时间间隔的人群活动流量的真实图和预测图进行热力图绘制,如图 11 所示。从预测结果上分析,NY 数据集的预测图与真实图更接近,SFO 对于流量高的区域预测结果较差;从人群活动区域分析,人群活动主要集中在固定的区域,说明人群活动遵循一定的空间规律;从时间语义上分析,不同时段同一区域的流量不同,比如凌晨的人群活动流量较其他时间间隔流量较少,说明人群活动遵循一定的时间规律。

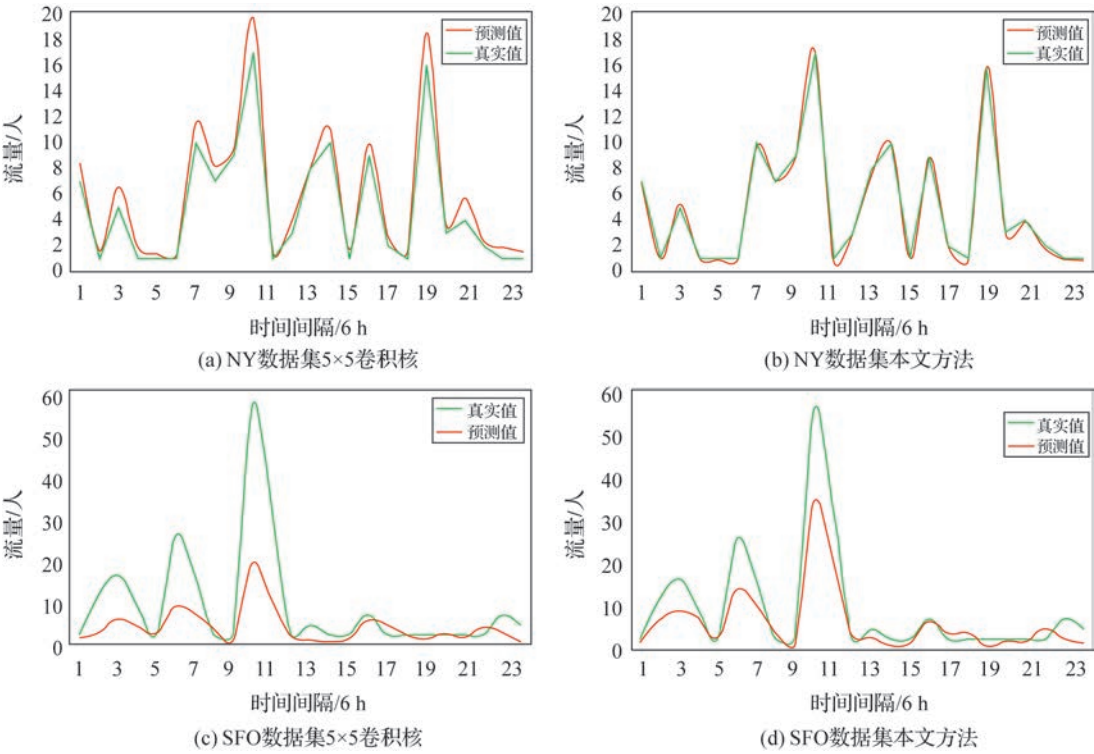


图 10 不同方案人群流量真实值与预测值拟合结果

Fig.10 Comparison of real value and predicted value of crowd flow in different schemes

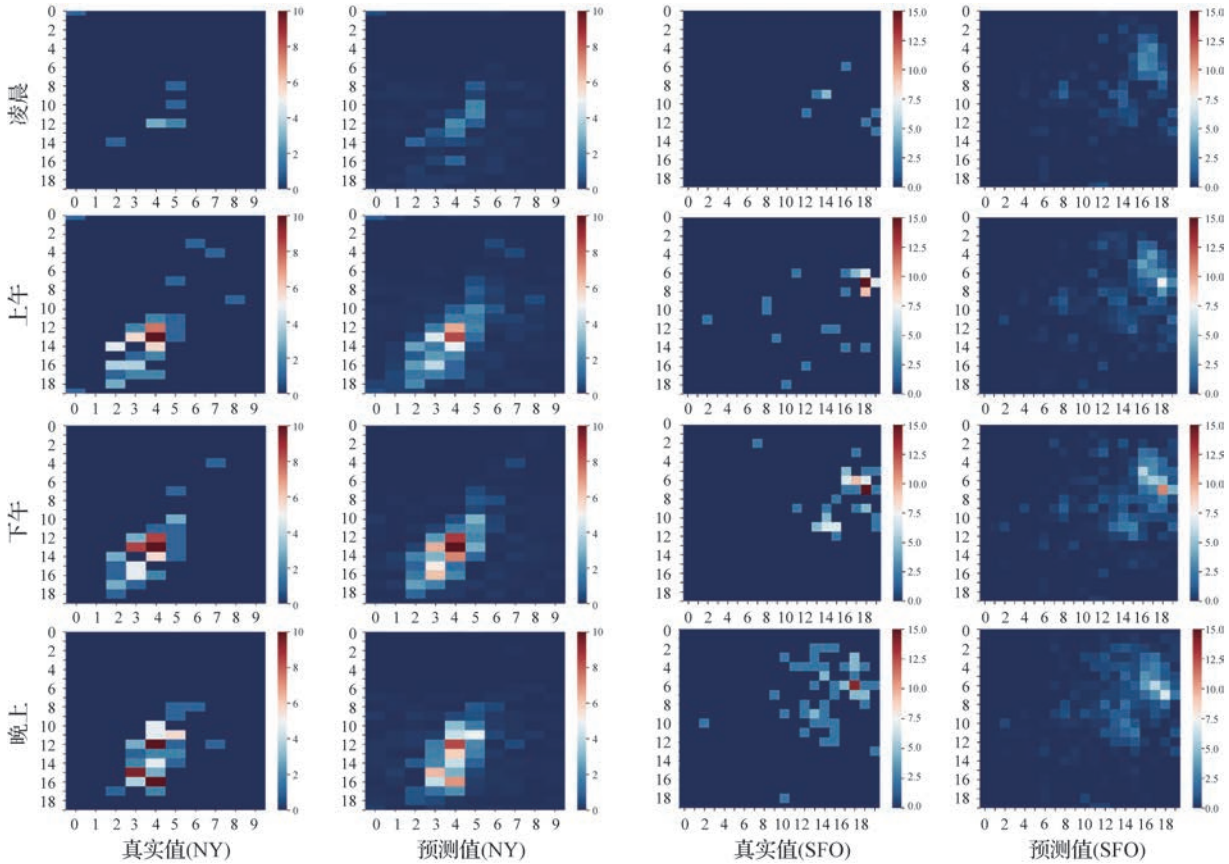


图 11 不同时间间隔人群流量真实值与预测值热力图

Fig.11 The visualization of real and predicted values of crowd flow at different time intervals

2.4 本文方法与常用时空预测方法精度对比

为验证本文方法 MST-Net 模型相比于其他方法的有效性,试验选取历史平均模型 HA^[9]、自回归平均移动模型 ARIMA^[10]、支持向量回归 SVR^[33]、门控循环单元 GRU^[29] 和时空动态网络 (spatial-temporal dynamic network, STDN)^[18] 5 种方法进行对比,结果见表 4。试验结果表明,在均方根误差和平均百分比误差上,MST-Net 模型优于其他方法,尤其在 NY 数据集上的表现更为显著,平均百分比误差最大降低了76.81%,表明了本文提出模型对人群活动流量预测任务上的有效性。传统的时间序列模型(HA 和 ARIMA),和基于回归的方法(SVR),在该任务上表现欠佳,主要因为这些方法一方面依赖于历史数据,忽略了空间因素,另一方面是对于这种具有周期性和趋势性的非平稳的时间序列数据难以很好的拟合;而 GRU 和 STDN 在该问题上取得了较好的效果,说明基于神经网络的方法能够较好地拟合复杂的非线性的时空数据。但是 GRU 在进行预测时,只提取了时间维度的特征,同样忽略了空间特征,STDN 虽然考虑了空间特征,却只考虑了单一尺度,而本文方法能够较好地拟合复杂的非线性时空关系且顾及空间多尺度特征,因而具有更高的精度。

表 4 本文方法与常用时空预测方法精度对比
Tab.4 Comparison with other spatio-temporal prediction methods

数据集	方法	均方根误差	平均百分比误差/(%)
NY	HA	1.255 9	79.73
	ARIMA	0.599 2	66.05
	SVR	1.214 5	73.91
	GRU	0.005 5	10.23
	STDN	0.003 7	3.18
	MST-Net	0.002 6	2.92
SFO	HA	0.607 6	78.56
	ARIMA	0.572 8	77.27
	SVR	0.582 2	83.94
	GRU	0.007 2	65.76
	STDN	0.001 2	35.73
	MST-Net	0.000 9	23.87

3 结 论

由于传统时空预测模型大多基于单一尺度的空间特征进行时空预测,难以顾及多尺度空间特征对预测结果的影响,因此,本文提出基于卷积神经网络且通过并联卷积方式融合空间多尺度特征的时空网络人群活动预测模型,并在两组真实的社交媒体签到数据

集上来验证模型的有效性。试验结果表明,与基于单一尺度空间特征的时空预测模型相比,基于多尺度空间特征的时空预测模型能够获取更完备的空间特征以提高预测精度,进而更好地提取人群活动在空间上的动态变化,实现人群活动规律的探索。

但是本文提出的方法是基于规则格网上的,受格网单元形状和大小的限制。一方面,时间尺度对于模型的边界效应和影响程度需要进一步试验和论证;另一方面,本文仅采用社交媒体签到数据,存在不能充分反映人群流量的问题。下一步可融合其他类型的感知数据,例如 GPS 轨迹数据等来反映人群流量;此外,人群活动还包含有丰富的空间语义信息,如何融合空间语义信息,实现人群活动规律的可解释性也是下一步研究的重点。

参考文献:

[1] ZHENG Yu. Trajectory data mining: an overview[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2015, 6(3): 29.

[2] MA Xiaolei, TAO Zhimin, WANG Yinai, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data[J]. Transportation Research Part C: Emerging Technologies, 2015, 54: 187-197.

[3] WANG Yingzi, ZHOU Xiao, NOULAS A, et al. Predicting the Spatio-temporal evolution of chronic diseases in population with human mobility data[C]// Proceedings of the 27th International Joint Conference on Artificial Intelligence. [S.l.]: IJCAI, 2018: 3578-3584.

[4] 王丽鲲. 基于社交媒体地理数据挖掘的游客时空行为分析[D]. 上海: 上海师范大学, 2017.

WANG Likun. Spatial temporal behavior analysis of tourists based on social media geographic data mining [D]. Shanghai: Shanghai Normal University, 2017.

[5] LIU Yu, LIU Xi, GAO Song, et al. Social sensing: a new approach to understanding our socioeconomic environments[J]. Annals of the Association of American Geographers, 2015, 105(3): 512-530.

[6] LI Xiaolong, GANG Pan, WU Zhaohui, et al. Prediction of urban human mobility using large-scale taxi traces and its applications[J]. Frontiers of Computer Science, 2012, 6(1): 111-121.

[7] VLAHOGIANNI E I, GOLIAS J C, KARLAFTIS M G. Short-term traffic forecasting: overview of objectives and methods[J]. Transport Reviews, 2004, 24(5): 533-557.

[8] OKUTANI I, STEPHANEDES Y J. Dynamic prediction of traffic volume through Kalman filtering theory[J]. Transportation Research Part B: Methodological, 1984, 18(1): 1-11.

[9] 刘静, 关伟. 交通流预测方法综述[J]. 公路交通科技, 2004, 21(3): 82-85.

LIU Jing, GUAN Wei. A summary of traffic flow forecasting methods[J]. Journal of Highway and Transportation Research and Development, 2004, 21(3): 82-85.

[10] DEVEAUD R, ALBAKOUR M D, MACDONALD C, et al.

- Experiments with a venue-centric model for personalised and time-aware venue suggestion[C]//Proceedings of the 24th ACM International on Conference on Information & Knowledge Management, Melbourne: ACM, 2015.
- [11] 董春娇, 邵春福, 诸葛承祥, 等. 拥挤流状态下城市快速路交通流时空特性[J]. 北京工业大学学报, 2012, 38(8): 1242-1246, 1268.
- DONG Chunjiao, SHAO Chunfu, ZHUGE Chengxiang, et al. Spatial and temporal characteristics for congested traffic on urban expressway[J]. Journal of Beijing Polytechnic University, 2012, 38(8): 1242-1246, 1268.
- [12] HU Wangsu, YAO Zijun, YANG Sen, et al. Discovering urban travel demands through dynamic zone correlation in location-based social networks[C]//Proceedings of 2018 Joint European Conference on Machine Learning & Knowledge Discovery in Databases, Dublin: Springer, 2018.
- [13] DENG Dingxiong, SHAHABI C, DEMIRYUREK U, et al. Latent space model for road networks to predict time-varying traffic[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data, California: ACM, 2016.
- [14] YANN L, YOSHUA B, HINTON G, et al. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [15] ZHANG Junbo, ZHENG Yu, QI Dekang. Deep Spatio-temporal residual networks for citywide crowd flows prediction[C]//Proceeding of the 31 AAAI Conference on Artificial Intelligence (AAAI-17), San Francisco: AAAI, 2016.
- [16] YAO Huaxiu, WU Fei, KE Jintao, et al. Deep multi-view spatial-temporal network for taxi demand prediction[C]//Proceedings of 2018 National Conference on Artificial Intelligence, Honolulu: AAAI, 2018: 2588-2595.
- [17] ZHAO Ling, SONG Yujiao, ZHANG Chao, et al. T-GCN: a temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(9): 3848-3858.
- [18] YAO Huaxiu, TANG Xianfeng, WEI Hua, et al. Revisiting spatial-temporal similarity: a deep learning framework for traffic prediction[C]//Proceedings of 2019 AAAI Conference on Artificial Intelligence, Palo Alto: AAAI, 2019.
- [19] BAHDANAU D, CHO K, BENGIO Y, et al. Neural machine translation by jointly learning to align and translate[C]//Proceedings of 2015 International Conference on Learning Representations, San Diego: ICLR, 2015.
- [20] 刘瑜. 基于空间大数据的社会感知[J]. 中国计算机学会通讯, 2015, 11(11): 27-32.
- LIU Yu. Social computing based on spatial data[J]. Communications of CCF, 2015, 11(11): 27-32.
- [21] 陈超, 齐峰. 卷积神经网络的发展及其在计算机视觉领域中的应用综述[J]. 计算机科学, 2019, 46(3): 63-73.
- CHEN Chao, QI Feng. Review on development of convolutional neural network and its application in computer vision[J]. Computer Science, 2019, 46(3): 63-73.
- [22] MOREIRA-MATIAS L, GAMA J, FERREIRA M, et al. Predicting taxi-passenger demand using streaming data[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(3): 1393-1402.
- [23] GOODCHILD M F. Models of scale and scales of modelling[M]//TATE N J, ATKINSON P M. Modelling Scale in Geographical Information Science, New York: John Wiley and Sons, 2001.
- [24] LUO Weijie, LI Yujia, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems, Red Hook: ACM, 2016.
- [25] XIA Chang, ZHANG Anai, WANG Haijun, et al. Modeling urban growth in a metropolitan area based on bidirectional flows, an improved gravitational field model, and partitioned cellular automata[J]. International Journal of Geographical Information Science, 2019, 33(5): 877-899.
- [26] 陆川伟, 孙群, 陈冰, 等. 车辆轨迹数据的道路学习提取法[J]. 测绘学报, 2020, 49(6): 692-702. DOI: 10.11947/j.AGCS.2020.20190305.
- LU Chuanwei, SUN Qun, CHEN Bing, et al. Road learning extraction method based on vehicle trajectory data[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(6): 692-702. DOI: 10.11947/j.AGCS.2020.20190305.
- [27] SCHMIDHUBER J. Deep learning in neural networks: an overview[J]. Neural Networks, 2015, 61: 85-117.
- [28] BENGIO Y, SIMARD P Y, FRASCONI P, et al. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157-166.
- [29] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [30] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[C]//Proceedings of 2014 Deep Learning and Representation Learning Workshop, Montreal: NIPS, 2014.
- [31] GAO Song, JANOWICZ K, COUCLELIS H. Extracting urban functional regions from points of interest and human activities on location-based social networks[J]. Transactions in GIS, 2017(21): 446-467.
- [32] 谢志文, 王海军, 张彬, 等. 城市扩展元胞自动机多结构卷积神经网络模型[J]. 测绘学报, 2020, 49(3): 375-385. DOI: 10.11947/j.AGCS.2020.20190147.
- XIE Zhiwen, WANG Haijun, ZHANG Bin, et al. Urban expansion cellular automata model based on multi-structures convolutional neural networks[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(3): 375-385. DOI: 10.11947/j.AGCS.2020.20190147.
- [33] SMOLA A J, SCHÖLKOPF B. A tutorial on support vector regression[J]. Statistics and Computing, 2004, 14(3): 199-222.

(责任编辑:张艳玲)

收稿日期: 2020-06-11

修回日期: 2021-01-15

第一作者简介: 李静(1990—),女,博士生,研究方向为时空数据挖掘和社会感知计算。

First author: LI Jing(1990—),female, PhD candidate, majors in spatio-temporal data mining and social concept computing.

E-mail: brandy12367@163.com

通信作者: 郭文月

Corresponding author: GUO Wenye

E-mail: guowyer@163.com