

# 融合多视图注意力机制的城市交通流量补全

康 雁 陈 铁 李 浩 杨 兵 张亚钊 卜荣景

云南大学软件学院 昆明 650500



**摘 要** 交通流量信息是智能交通系统和城市计算的重要基础。交通流量数据作为新型时序数据,由于数据的采集方式和外部复杂因素的影响,使得数据缺失现象是常见且无法避免的。如何有效地挖掘交通流量数据的时空特性和数据间的关联成为了提高缺失数据补全精度的关键。传统的统计学方法不能满足日益增长的数据需求,深度学习的应用推动了缺失数据的补全方法向更高的精确度发展。文中深入分析了交通流量的时间特性和空间分布,对交通流量的缺失情况进行了假设,提出了一种UMAtNet(U-net with Multi-View Attention Mechanisms)交通流量补全模型。该模型将短期的、趋势的、周期的时间数据与空间数据融合,同时采用不同的数据相关性测量方法,融合了一种多视图注意力机制,能够优化模型对缺失部分数据空间相关性的影响。为了验证模型的有效性,文中使用北京交通轨迹开源数据集进行实验,并在实验中详细地分析了模型各部分和损失函数对补全精度的影响,实验结果表明,UMAtNet和相应组件融合能进一步提高补全精度。

**关键词:** 交通流量补全;时空特征;多视图;U-Net;注意力机制

**中图法分类号** TP311

## Urban Traffic Flow Completion with Multi-view Attention Mechanism

KANG Yan, CHEN Tie, LI Hao, YANG Bing, ZHANG Ya-chuan and BU Rong-jing

School of Software, Yunnan University, Kunming 650500, China

**Abstract** Traffic flow information is an important basis for intelligent transportation systems and urban computing. Traffic flow data is a new type of time series data. Due to the data collection method and the influence of external complex factors, the phenomenon of data loss is common and unavoidable. How to effectively mine the spatial-temporal characteristics of traffic flow data and the correlation between the data becomes the key to improve the missing data completion accuracy. Traditional statistical methods cannot meet the increasingly complex data requirements, and the application of deep learning promotes the development of missing data completion methods to higher accuracy. The article deeply analyzes the spatial-temporal characteristics of traffic flow, makes assumptions about the missing traffic flow, and proposes a UMAtNet (U-net with Multi-view Attention Mechanisms) traffic flow complement model. The model fuses closeness, trend and period time data with spatial data, and adopts different data correlation measurement methods to fuse a multi-view attention mechanism, which can optimize the impact of the model on the spatial correlation of missing data. In order to verify the model, we use the open source data set of Beijing traffic data in the experiment, and analyzes in detail the influence of each part of the model and the loss function on the completion accuracy. The experimental results show that the fusion of UMAtNet and corresponding components can further improve the completion accuracy.

**Keywords** Completion of traffic flow, Spatial-temporal characteristics, Multi-view, U-Net, Attention mechanism

## 1 引言

随着交通传感器的普及和交通传感器技术的发展,交通数据呈爆炸式增长,准确和及时的交通流量信息对于智能交通系统(Intelligent Transportation Systems, ITS)的成功部署十分重要,它能够帮助交通管理部门做出更好的出行决策,从而缓解交通拥堵,减少碳排放,提高交通运营效率<sup>[1-2]</sup>。其中,

交通流量数据是城市计算和智能交通系统的一个重要的数据基础,交通流量数据很大程度上依赖于从各种传感器来源收集的历史和实时交通数据,包括雷达、摄像头、移动全球定位系统、社交媒体等。然而,数据采集和统计过程中往往存在数据缺失的情况,这是一个常见且难以避免的问题。

造成数据缺失的原因是多方面的,主要有:1)部分数据遗漏和丢失,当我们在一个交通基础设施建设不完善、交通数据

到稿日期:2020-08-12 返修日期:2020-11-11 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:基于迁移学习的动态多目标需求优选研究(61762092);云南省重大科技专项(202002AB080001)

This work was supported by the Research on Dynamic Multi-objective Requirement Optimization based on Transfer Learning(61762092) and Major Science and Technology Project in Yunnan Province(202002AB080001).

通信作者:康雁(kangyan@ynu.edu.cn)

采集技术尚不成熟或刚刚兴起的城市进行流量统计时,由于数据采集设备、存储介质或传输媒体的故障,以及一些人为因素等导致数据缺失;2)部分数据暂时无法获取,目前大多数交通流量研究采用出租车轨迹数据,出租车轨迹数据易于获取和统计,而由于交通工具的多样性,人们依靠自行车或者私家车出行的数据通常因个人利益和隐私问题而无法获取。从数据来源上来说,流量数据的统计是不完整的。因此,我们采用了不同的方式模拟了数据的缺失情况,并将缺失的数据作为我们研究问题的输入,具体内容见 3.1 节。

流量数据丢失的问题降低了许多应用程序的性能。例如,由于数据采集和统计存在缺失的情况,以及交通过程的复杂性,交通预测模型往往效率较低,预测问题变得更具挑战性。综合上述内容,本文的主要贡献有:1)借助图像补全的方法,提出了一种基于时空特性的城市交通流量补全模型,能够对缺失数据实现准确快速的建模;2)分析了数据的特性和空间分布之后,基于不同的数据相关性测量方法,在补全模型中融合了改进的多头注意力机制,得到区域交通流量的多个特征表示,这些特征从不同视角揭示了空间中区域流量的不同特性,称为多视图特征,该方法优化了模型对缺失部分数据空间相关性的影响;3)采用了一种跳跃连接的方式,使本文模型在考虑缺失区域与相近区域的相互影响上,更加侧重于缺失数据的补全;4)融合了外部数据的影响,如天气、节假日信息,并将该部分融入到缺失流量的补全模型。在实验中,我们详细地分析了模型的各个模块对实验任务的影响,并对部分内容进行了可视化,同时将本文提出的模型和多种模型进行了比较,实验结果表明本文方法是有效的。

## 2 相关工作

### 2.1 基于统计学方法的交通流量补全

缺失数据的问题广泛存在于许多交通信息系统中,这给进一步的研究带来了很大的麻烦。为了解决这个问题,较早的方法是直接基于历史数据的估算方法,即基于相同时间间隔的历史数据的平均值来填补位置缺失的数据点<sup>[3-5]</sup>。此外,还有最大值补全法、中值补全法、时间序列补全法、线性回归补全法(Linear Regression, LR)等<sup>[6-7]</sup>。其中,均值补全法、最大值补全法和中值补全法是利用历史流量数据的均值、最大值或中值作为当前缺失数据的插补值,这些方法有一定的优点,如简单易行、方便快捷处理数据和补全。但是,这些方法简单地利用日周期性、周周期性等时间特征来进行补全,无法反映真实的数据变化情况,也不能消除空间相邻区域的影响。时间序列补全法可以较好地利用时序流量数据建立模型,对缺失的短时交通流量进行补全。线性回归可以分析影响缺失流量的因素,建立自变量和因变量之间的线性关系,预测缺失的流量数据。然而,这些方法没有进一步挖掘交通流量随时间变化的特性,以及数据之间的关联性,随着数据缺失率的增加,补全精度不高。Tan 等<sup>[8]</sup>提出了一种基于张量分解的数据插补方法,该方法继承了基于矩阵模型的缺失值补全的优势,挖掘了交通数据的多维内在相关性。Ran 等<sup>[9]</sup>同样基于张量的方法对时空缺失交通数据进行补全,融合不同时间段区域的流量数据,充分利用临近位置的空间信息,对缺失的位

置进行补全,该方法适用于一些数据长时间丢失的情况。但是,这两种方法对具体时刻的缺失数据和短时间内交通流量的缺失缺少明确的补全方案,忽略了短时间内数据的相关性。综上,基于统计学的方法不能系统地针对整个城市的交通流量进行补全,没有深度挖掘数据的时空特性,补全的精度有待提高。

### 2.2 基于机器学习的交通流量补全

近年来,利用机器学习算法补全缺失交通流量数据逐渐被应用。对于缺失交通流量数据补全的研究开始考虑数据的时间依赖性和空间相关性。Tian 等<sup>[10]</sup>提出了一种基于长短期记忆(Long Short-Term Memory, LSTM)的新颖方法,采用多尺度时间平滑来推断丢失的数据,但是该方法需要较长的数据时间跨度才能达到良好的效果。Ku 等<sup>[11]</sup>采用 K-means 聚类技术将交通流量相似的路段组合在一起,然后为每路段构建基于堆叠自动编码器的深度学习模型,提取其时空关系,并将所提取的时空关系用于估算缺失的数据。但是,该方法不能针对整个城市的流量数据进行补全,且效率较低。Yu 等<sup>[12]</sup>基于三维卷积网络(3D Convolutional Neural Networks, Conv3D),提出了一种称为 TF-3DNet 的方法,以实现准确地预测大规模的交通流量,其关键思想是使用 3D 卷积核同时提取和融合时空交通流量数据中的特征。Li 等<sup>[13]</sup>以生成式对抗网络(Generative Adversarial Network, GAN)模型为基础,融合三维卷积进行交通流量的补全。Wang 等<sup>[14]</sup>以 GAN 作为交通流量数据补全方法,以路段实际流量为基础,利用图像化的方式生成路网二维信息图,实现路网二维信息图缺失部分的补全。Yuan 等<sup>[15]</sup>在常规的缺失数据补全方法的基础上利用 DCGAN 模型的思想,加入补全损失函数和判别损失函数,利用数据生成思想进行区域交通流量数据的补全。以上方法以 GAN 模型和 Conv3D 模型为基础,部分进行了交叉融合,但是没有进一步分析数据本身的关联性和模型对缺失数据补全效果的重要性,因此补全精度不高。

### 2.3 基于 U-Net 的图像补全

目前,深度学习方法在图像的补全和修复方面取得了良好的效果。Hong 等<sup>[16]</sup>提出了一种基于 U-Net 的图像修复方法,该方法可以分为粗修和精修两步,其中粗修复网络和精细修复网络整体上是 U-Net 架构,并在结构上做了调整和优化,达到了对数据缺失部分的有效预测。Guilin 等<sup>[17]</sup>提出了一种部分卷积(Partial Convolution, Pconv),其中将卷积屏蔽并重新规范化为仅以图像有效像素为条件,并应用一系列损失函数来衡量模型对图像的补全效果。Wen 等<sup>[18]</sup>在点云补全中使用了一种 Skip-attention 机制,该机制直接搜索输入飞机形状中的信息局部区域,并以此来补全缺失图像。以上方法在图像补全方面取得了良好效果。和上述方法相近的是,我们采用的缺失流量数据以流量图的形式呈现。借助图像补全的部分先进性方法,我们对模型进行了改进。对于交通流量补全任务,我们针对交通流量的特殊性,搭建了一种基于时空特性的交通流量补全模型,同时融合了自身注意力和多头注意力机制的优点,并且使用了一种更直接的跳跃连接的方式(skip-connection),能够针对性地对缺失部分数据进行补全,从而提高补全的精确度。我们改进的模型和方法在

补全任务上取得了良好的效果,超越了现有的流量补全方法。

### 3 融合多视图注意力机制的补全模型

本文模型可以有效地进行交通流量的补全,如图 1 所示。本文设计了一种融合多视图注意力机制的补全模型(UMAtNet),该模型在 U-net<sup>[19]</sup>的基础上进行了改进,并融入了基于多视图特征的注意力机制,具体细节见 3.2.2 节。该模型能够根据历史时刻的交通流量数据,结合当前时刻的未缺失数据对缺失数据进行补全,以达到更加精确的补全效果。我们在模型中加入了跳跃连接,使模型侧重于对数据缺失部分的补全效果,这一点在计算损失函数部分时十分重要。本文通过实验对上述方法进行了对比,实验细节见第 4 节。

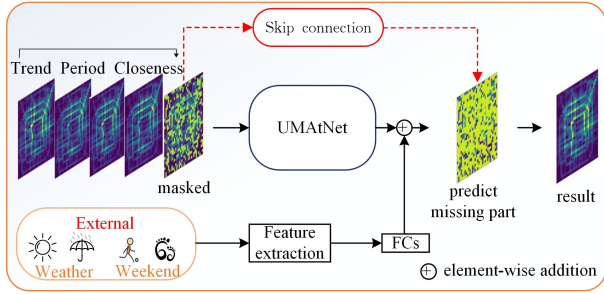


图 1 补全模型结构图

Fig. 1 Completion model structure

本节首先对问题进行定义并介绍相应的掩模机制和模型输入,假设了数据的缺失情况,然后进一步讨论补全模型的各个模块和采用的先进性方法。

#### 3.1 问题定义与假设

**定义 1** 一个城市被划分为  $H \times W$  个相同大小的网格(如  $1\text{ km} \times 1\text{ km}$ )。其中,  $H \times W$  为  $32 \times 32$ , 每个网格称为区域<sup>[20]</sup>, 每个区域包含的交通流量时间间隔为 0.5 h, 即交通流量每 0.5 h 统计一次, 得到每日的数据量为  $T$ , 这里  $T$  为 48。对于每一个区域, 有两种交通流量类型, 即流入和流出, 因此城市的日流量可表示为  $X \in \mathbf{R}^{T \times H \times W \times 2}$ 。

**问题假设:** 通过对交通流量数据的分析, 我们注意到流量数据会在不同的时刻发生不同程度或形式的丢失。因此, 对于不同时刻的流量图(不同样本), 我们采用有差异的二进制掩码去处理此类情况。我们设置了两种不同的二进制掩码, 如图 2 所示, 一种是缺失率  $P$  不同的二进制掩码,  $P=10\%$  表示缺失数据占整体数据的 10%, 图 2 给出了不同缺失率下呈现的结果。

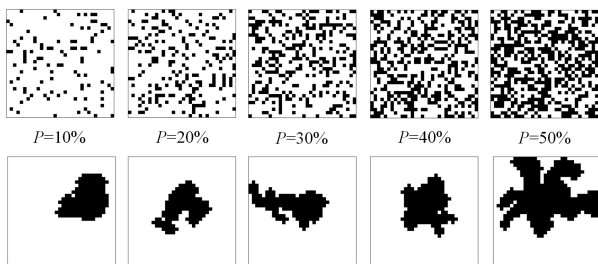


图 2 不同缺失率的掩模机制以及不规则的掩模机制

Fig. 2 Mask mechanisms with different missing rates and irregular mask mechanisms

另一种是不均匀分布的二进制掩码, 即模拟数据成块缺失的情况。其中掩码的大小与输入的流量图大小和通道数一致, 以上两种掩模机制均表示为  $X_{\text{mask}} \in \mathbf{R}^{H \times W \times C}$ ,  $x_{ij} = 0$  表示缺失, 其中  $C=2$  表示对同一时刻的入流量图与出流量图采用相同的掩码机制。然后, 对于同一时刻的两种流量矩阵, 与二进制掩码矩阵相乘, 如图 1 所示, 得到叠加到交通流量上的结果  $X_{\text{masked}} \in \mathbf{R}^{H \times W \times 2}$ , 假设当前时刻( $t$ )流量数据的缺失, 同时保留了未缺失部分的数据, 因此我们将其作为模型的主要输入。其中, 对于不同的数据样本, 我们采用了随机的掩模机制, 即每个样本的缺失情况是不同的。

**定义 2** 基于不同的时间角度, 为了有效利用历史时刻的交通流量数据, 我们将与缺失数据时刻相关的邻近性数据  $X_{t-1}$ 、周期性数据  $X_{t-T}$ 、趋势性数据<sup>[21]</sup>  $X_{t-T * n}$  和当前时刻( $t$ )数据  $X_{\text{masked}}$  进行融合, 作为模型的第一部分输入。第二部分输入为外部因素, 如天气、节假日数据等, 表示为  $X_{\text{ext}}$ , 考虑了外部因素对模型补全效果的影响。综上, 本文模型输出可表示为:

$$X_{\text{out}} = f(X_{\text{ext}}; \text{Concat}(X_{t-T * n}, \dots, X_{t-T}, X_{t-1}, X_{\text{masked}})) \quad (1)$$

其中,  $f$  为我们采用的模型和相应的参数,  $X_{\text{out}}$  为待补全时刻的模型输出。而对于我们的补全任务, 我们的目标是得到模型输出  $X_{\text{out}}$  中对缺失数据的预测值, 因此我们在模型损失计算上进行了修改, 得到了最终的模型输出  $\hat{X}_{\text{result}}$ , 具体细节见后文的 3.3 节及式(12)。

#### 3.2 UMatNet 补全模型架构

考虑到 U-Net 模型在图像补全领域和图像分割领域取得了良好效果<sup>[17-19]</sup>, 本文模型以 U-Net 为基础, U-Net 的 U 形结构如图 3 所示。该网络是一个全卷积网络, 网络的输入是一张融合时序数据的流量图, 即式(1)中时序数据由远及近连接的结果, 该流量图为  $X_{\text{in}} \in \mathbf{R}^{32 \times 32 \times C}$ , 其中  $C$  为我们融合的时序出入流量通道。由于我们的交通流量图是低分辨率的, 因此我们仅进行了 4 次下采样和 4 次上采样, 每次采样过程使用了一个上采样(下采样)和两个卷积(卷积核  $3 \times 3$ )层, 与 Olaf 等<sup>[19]</sup>提出的方法不同的是, 我们将上采样过程的结果进行了融合, 即保存了模型对缺失部分数据的重建过程, 最后将模型上采样融合的结果作为本文注意力机制的输入, 通过多视图注意力层得到本文模型的输出结果  $X_{\text{out}} \in \mathbf{R}^{32 \times 32 \times 2}$ 。

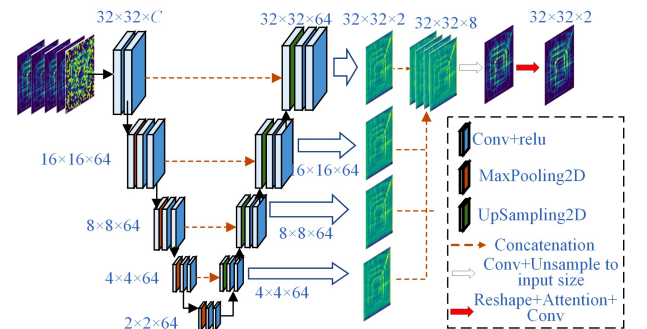


图 3 融合多视图注意力的 U-Net 模型

Fig. 3 U-Net with multi-view attention mechanisms(UMAtNet)

### 3.2.1 数据相关性特征

根据定义 1, 我们以历史流量的日平均值作为计算区域数据相似性或相关性的依据, 如图 4—图 6 所示, 以入流量为例, 展示了不同数据相关性计算方式下的日交通流模式。区域的日交通流量表示为  $Y_N = \{y_{i,t} | i = 1, \dots, N, t = 1, \dots, T\}$ , 其中  $N = H \times W$ , 表示某一时刻有  $N$  个区域, 每个区域的日交通数据量为  $T$ 。由于数据量较大, 图中只显示了部分区域的相似性数据。两个不同的区域  $X$  和  $Y$  可表示为:

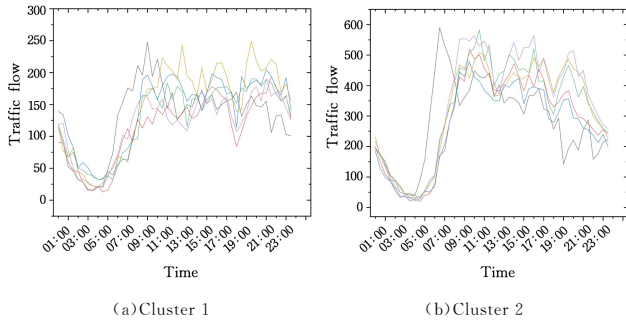
$$X = \{x_1, x_2, \dots, x_T\}$$

$$Y = \{y_1, y_2, \dots, y_T\}$$

我们采用均方根误差 (Root Mean Square Error, RMSE) 的计算方式, 该方式作为一种常用的测量数值之间差异的度量, 首先根据 RMSE 衡量区域数据间的相似性, 依次计算出两个不同区域  $X, Y$  的数据相似度如下:

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (x_i - y_i)^2} \quad (2)$$

图 4 给出了某个地区及其相似性区域的日流量模式, 可以看出, 以 RMSE 为度量方式, 各区域的出行模式较为相似且数值相近, 在流量数据上表现出了良好的数据相似性。但是, 由于该方法是一种较为直接和简洁的计算方式, 不能深入地分析时序交通流量的特性。受 Wang 等<sup>[21]</sup>、Vu 等<sup>[22]</sup>、Tang 等<sup>[23]</sup>研究的启发, 我们从另外两个角度分析了交通流量数据的相关性及其特性。



注: 区域的出行模式较为相似, 且数值相近

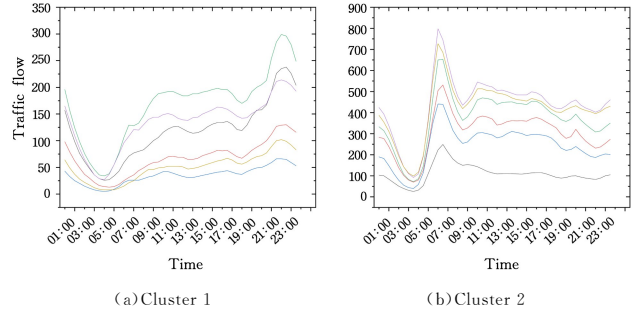
图 4 使用均方根误差衡量相似区域流量

Fig. 4 Using RMSE to measure traffic flow in similar areas

Wang 等<sup>[21]</sup>使用皮尔逊相关系数来测量区域之间的人群登记相似性, 在自行车数据上进行分析, 验证了该方法对城市间相似区域匹配的有效性。为了进一步采用皮尔逊相关系数来衡量城市内交通流量的相关性, 我们以此为依据, 计算两个不同区域的皮尔逊相关系数  $\rho = \text{Pearson}(X, Y)$ ,  $i \neq j$  and  $i, j \in N$ 。具体计算式如下:

$$\rho = \frac{[\sum_{i=1}^T (x_i - \bar{X})(y_i - \bar{Y})]^2}{\sum_{i=1}^T (x_i - \bar{X})^2 \sum_{i=1}^T (y_i - \bar{Y})^2} \quad (3)$$

相关系数的绝对值越大, 相关性就越强, 我们以  $\rho > 0.8$  为例, 如图 5 所示, 可以看出, 使用皮尔逊相关系数作为衡量标准的相似性区域, 交通流量在趋势上较为相似, 如呈现出不同的出行早高峰和晚高峰, 但交通流量在数值上可能相差较大, 即不同区域同一时刻的人群流量相差较大。



注: 对区域的出行模式衡量较好, 但数值差异可能较大

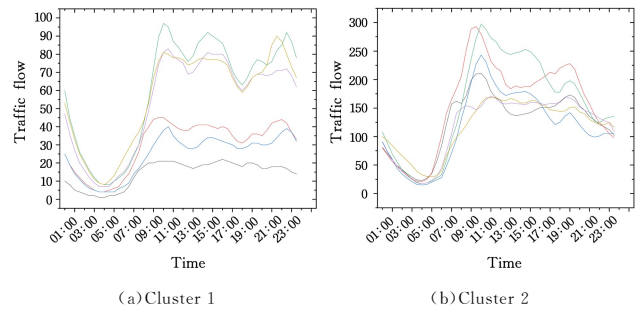
图 5 使用皮尔逊相关系数衡量相似区域的流量

Fig. 5 Using Pearson correlation coefficient to measure traffic flow in similar areas

动态时间规整 (Dynamic Time Warping, DTW) 是语音识别中较为经典的一种算法, Vu 等<sup>[22]</sup>在语音识别任务中得到动态时间规整与相关系数之间的组合的调查结果, 取得了很高的识别准确性。Tang 等<sup>[23]</sup>使用基于 DTW 的分解探针数据开发了一种非参数模型, 将数据分为不同的信号相位, 而不是识别确切的信号模式, 可反映交通信号周期中行驶时间的变化。因此, 考虑到交通流量的时序性, 以及其在短时间内的变化, 我们还使用了 DTW 算法对区域进行了相似性的计算。为了比较这两个序列的出行模式, 我们需要构造一个  $N \times N$  的矩阵网格  $D$ , 矩阵元素  $D(i, j)$  表示  $x_i$  和  $y_j$  两个点的距离  $d(x_i, y_j)$ , 距离越小则相似度越高, 这里使用的是欧氏距离。计算  $D$  的递归方程可表示为:

$$\begin{cases} D(1, 1) = d(x_1, y_1) \\ D(i, j) = d(x_i, y_j) + \min\{D(x_{i+1}, y_j), \\ D(x_i, y_{j+1}), D(x_{i+1}, y_{j+1})\} \end{cases} \quad (4)$$

如图 6 所示, 使用 DTW 进行区域相似性的计算, 可以看出不同区域的早高峰和晚高峰出现在不同的时段, 如某区域的早高峰在 9:30, 但是其相似区域的早高峰可能在 8:30 或者 10:00, 这同样也表明交通流量数据可能存在滞后性。与之前相似性区域的计算方法不同的是, 距离计算的时间段是动态的。基于该方法, 我们可以挖掘出不同区域在较短时间内的相似性, 从而减小流量数据的滞后性带来的影响。需要注意的是, 使用 DTW 算法耗时较长。



注: 相似区域的早高峰和晚高峰出现在不同的时段

图 6 使用动态时间规整衡量相似区域流量

Fig. 6 Using DTW to measure traffic flow in similar areas

最后, 我们融合了上述 3 种交通流量相似性度量方法, 将城市范围的所有区域  $N = H \times W$  进行了相似性的计算, 如

图 7 所示。图 7 给出了不同相似性计算方式的结果,每一种方法的计算结果均表示为  $W_{CAi} \in \mathbf{R}^{N \times N}$ , 其中  $W_{(i,j)}$ ,  $i, j \in N$  表示第  $i$  个区域与第  $j$  个区域的相似性系数。

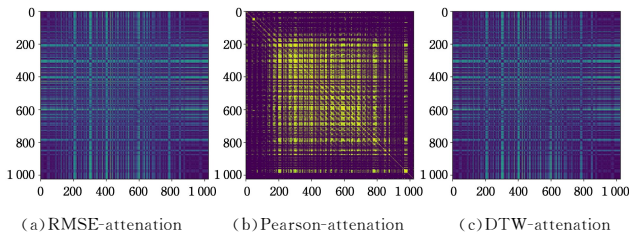


图 7 不同的相关性注意力矩阵

Fig. 7 Different correlation attention matrix (CA<sub>i</sub>)

### 3.2.2 多视图注意力机制

为了有效利用数据的空间特性,受多视图特征和注意力机制的启发,我们在多头注意力机制中融合了数据相似性的结果。其中,多头注意力机制建立在按比例缩放的点积注意力的基础之上<sup>[24-25]</sup>,如图 8 所示。该注意力作用于查询  $Q$ 、键  $K$  和值  $V$ ,具体为:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{Q * K^T}{\sqrt{d_k}}\right) * V \quad (5)$$

其中,  $d_k$  比例缩放因子,该因子的融入可以使点积计算结果更稳定,同时能加快模型的收敛速度。在融合了上述数据相关性注意力矩阵  $W_{CAi}$  后,我们得到了一种相关性点积注意力,具体为:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{Q * K^T}{\sqrt{d_k}}\right) \circ W_{CAi} * V \quad (6)$$

其中,  $\circ$  表示哈达玛积 (Hadamard product), 即矩阵对应元素相乘;  $*$  表示矩阵乘法, 即点积。

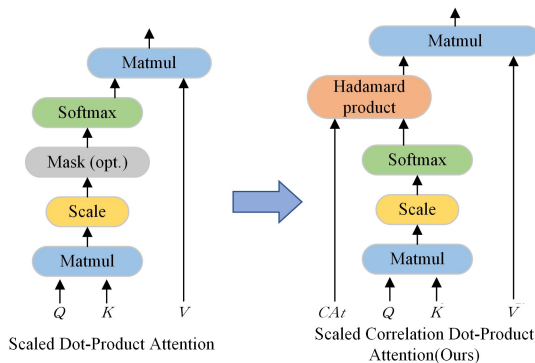


图 8 改进的点积注意力

Fig. 8 Improved dot product attention

然后,我们建立了一种多视图注意力机制,以多头注意力的方式来实现,获得了  $(Q, K, V)$  的  $h$  个不同的表示形式,如图 9 所示。计算了多视图相关性点积注意力之后,将结果连接起来,并用前馈神经网络层连接,使用与式 (6) 相同的计算方法来表示:

$$Head_i = Attention(Q W_i^Q, K W_i^K, V W_i^V) \quad (7)$$

$$MultiHead(Q, K, V) = \text{Concat}(Head_i) W^O \quad (8)$$

其中,我们使用了 3 种数据相关性度量方法,即  $h=3, i \in \{1, 2, 3\}$ ,  $W_i^Q, W_i^K, W_i^V, W^O$  是模型学习到的参数。这样做的优势在于,原来的多头注意力仅建立在按比例缩放的点积注意力

之上,仅由当前模型的输入输出来学习注意力矩阵,即当前时刻的数据空间相关性,而我们的多视图注意力机制建立在改进过后的数据相关性点积注意力之上,能够从不同的角度得到数据的特征表示,即模型能够得到空间区域的多视图特征表示,同时包含了历史所有数据的空间相关性以多视图注意力呈现,因此学习到的注意力更为稳定。对于数据的补全来说,考虑采用的相关性区域也更为精确。针对这一点,本文在对比实验中作了比较,结果表明本文方法是有效的。

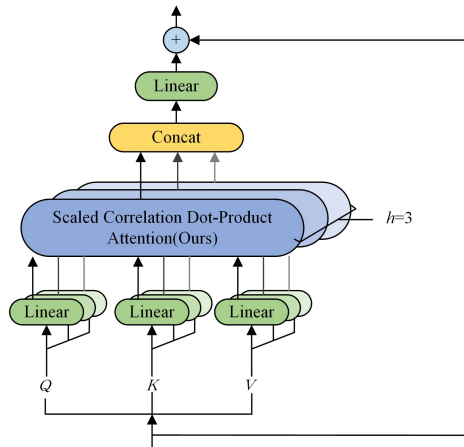


图 9 多视图相关性注意力

Fig. 9 Multi-view correlation attention (MAT)

### 3.3 跳跃连接

本文介绍的上述补全模型考虑了周边区域的影响,但由于补全任务的特殊性以及掩码机制的设定,本文模型应当侧重于每个缺失的区域的补全。

Guilin 等<sup>[17]</sup>应用了一系列损失函数,用于提升像素重建的精度,其中  $M$  为二进制掩码,用于区分数据的缺失和未缺失部分,然后以  $X_{in}$  为输入,经过模型得到预测值  $X_{out}$ ,以及待预测图像的真实值  $X_{gr}$ 。首先定义了缺失部分的补全损失:

$$L_{\text{masked}} = \frac{1}{N} \| (1 - M) \circ (X_{out} - X_{gr}) \|_1 \quad (9)$$

其中,  $L_{\text{masked}}$  表示模型预测值中缺失部分与对应的地面真实值之间的差距,  $\| \dots \|_1$  代表 L1loss,而且还定义了  $L_{\text{valid}}$ 。

$$L_{\text{valid}} = \frac{1}{N} \| M \circ (X_{out} - X_{gr}) \|_1 \quad (10)$$

其中,  $L_{\text{valid}}$  表示模型预测值中未缺失部分与对应的实际地面真实值之间的差距。由于图像补全任务与交通流量补全任务不同,本文在 Guilin 等<sup>[17]</sup>提出的损失函数上做了调整,没有使用实验中对本文补全任务影响较小的损失函数部分,而借鉴了上述两个损失函数,将其融合在一起并作为本文的对比损失函数,可表示为:

$$L_{\text{total}} = 6 L_{\text{masked}} + L_{\text{valid}} \quad (11)$$

其中,损失项权重由多次实验得出。对于本文的交通流量补全任务,我们对比了上述损失函数的计算方法,与 Guilin 等<sup>[17]</sup>不同的是,我们修改了计算损失函数的组成部分,使用一种跳跃连接 (skip-connection) 的方式,可表示为:

$$\hat{X}_{\text{result}} = (1 - M) \circ X_{out} + M \circ X_{gr} \quad (12)$$

其中,  $M \circ X_{gr}$  为我们用于跳跃连接的部分,代表原数据中的真

实值,即输入中未缺失部分的值;模型的最终输出为 $\hat{X}_{\text{result}}$ 。与依靠全局空间信息来完成交通流量补全相比,我们的跳跃连接直接搜索输入中缺失部分区域的局部信息,并有选择地使用这些区域来预测缺失部分的值,以达到更具有针对性的目的。实验结果表明,本文的跳跃连接是有效的,详见 4.4.1 节。

### 3.4 损失函数

本文模型的损失函数为:

$$MSE = \frac{1}{H \times W} \sum_{i=1}^{H \times W} (X_i - \hat{X}_i)^2 \quad (13)$$

其中, $X_i$ 表示每个区域的交通流量的真实值, $\hat{X}_i$ 表示每个区域的交通流量的预测值。我们在对比实验中还使用了 MAE 以及式(11)表示的 $L_{\text{total}}$ ,具体实验结果和细节见 4.4.2 节。

$$MAE = \frac{1}{H \times W} \sum_{i=1}^{H \times W} |X_i - \hat{X}_i| \quad (14)$$

## 4 实验

本节首先介绍了本文使用的北京市出租车轨迹数据集<sup>[20]</sup>,然后介绍了与其他算法的对比实验,最后介绍了不同的实验设置对实验结果的影响。

### 4.1 数据集介绍

北京的出租车轨迹数据和天气数据包括 4 个时间段。2013 年 6 月 1 日至 2013 年 8 月 30 日,2014 年 5 月 1 日至 2014 年 6 月 30 日,2015 年 5 月 1 日至 2015 年 6 月 30 日,2015 年 12 月 1 日至 2016 年 4 月 10 日。数据的时间间隔为 30 min,即 0.5 h 统计一次,有效的样本总量为 22459。根据定义 1 和定义 2,本文对所有的样本都进行了数据缺失的假设,选择最后 4 个星期的数据作为测试集,其他数据作为训练集,其中验证集占训练集的 10%。另外,还有对应的天气和节假日数据,天气数据包括 16 种数据类型,以 one-hot 编码的形式存储。

### 4.2 实验环境及设置

本文使用的实验环境及部分参数设置如表 1 所列。具体参数设置如下:模型中的激活函数为 relu,优化器为 adam,学习率为 0.0002,数据缺失率 Missing-rate  $P$  设置为 0.1~0.5。

表 1 实验环境及设置

Table 1 Experimental environment and settings

OS	Ubuntu 18.04.3 LTS
Memory	128 GB
CPU	Intel® Xeon(R) Gold 6130 CPU@ 2.10 GHz × 64
GPU	GTX1080Ti
Number of GPU cards	4
CUDA version	9.0
cuDNN version	7.3.1
tensorflow version	1.5.0
Keras version	2.1.6

### 4.3 实验结果

为了衡量模型的有效性,本文采用的评价指标是均方根误差(Root Mean Square Error, RMSE),其具体定义为:

$$RMSE = \sqrt{\frac{1}{H \times W} \sum_{i=1}^{H \times W} (X_i - \hat{X}_i)^2} \quad (15)$$

其中, $X_i$ 和 $\hat{X}_i$ 分别代表真实值和预测值。

本文将 UMAAtNet 与 9 种方法进行比较。

(1)HA。通过计算缺失区域历史时刻的出流量和入流量的均值来补全缺失时刻的出入流量数据。

(2)Linear Regression。一种常见的机器学习回归算法,根据历史流量数据建立缺失值和真实值的因果关系,对目标时刻缺失流量数据进行补全。

(3)AdaBoost Regression。一种统计学习回归算法,根据历史流量数据迭代训练机器学习器来进行补全。

(4)DCGAN<sup>[14-15]</sup>。以生成对抗网络为基础,能够通过生成器进行数据补全,通过判别器提升补全效果。

(5)ConvLSTM2D<sup>[26]</sup>。一种基于 LSTM 的网络,特征变换通过卷积来实现,能够同时捕捉流量数据的长时特性,同时提取空间特征。

(6)Conv3D<sup>[12-13]</sup>。一种三维卷积网络,将数据处理为类似视频数据的类型,能够较好地提取和利用数据时序信息。

(7)TCN<sup>[27]</sup>。一种时间卷积网络,注重数据的时序特征提取,但是数据需要进行特征变换,缺少对空间特性的捕捉。

(8)PconvNet<sup>[17]</sup>。一种图像补全领域的深度学习模型,我们对其进行了部分调整,用于交通流量的补全。

(9)U-Net<sup>[16-19]</sup>。一种用于图像分割和图像补全的模型,能够对历史流量数据进行建模,以达到较好的数据补全效果。

从表 2 可以看出,随着数据缺失率的增加,各模型的补全精度均有所下降。其中,历史均值插补由于较为简单,数据整体的补全精度较低。线性回归等统计学方法相比均值插补方法,在不同缺失率的情况下补全效果比较稳定,但同样达不到较高的补全精度。一些常用的深度学习模型的补全效果良好,但模型整体效果还有待提升。

表 2 各模型实验结果(RMSE)的对比

Table 2 Comparison of experimental results (RMSE) with base line

Model/Missing-rate $P$	10%	20%	30%	40%	50%
HA	14.74	20.64	25.20	29.19	32.70
Linear Regression	17.14	18.75	19.86	20.77	21.66
AdaBoost	13.58	18.85	23.26	26.51	30.10
DCGAN	5.22	7.16	9.75	10.81	12.83
ConvLSTM2D	8.60	9.02	13.01	16.83	17.89
Conv3D	5.70	6.51	7.24	10.33	11.92
TCN	13.71	14.10	14.30	14.65	15.08
PconvNet	17.71	18.77	20.70	21.58	22.05
U-Net	5.60	6.76	7.54	9.22	12.14
UMAAtNet	3.80	5.10	5.98	6.79	7.78

我们将部分模型的补全结果进行了可视化,如图 10 所示。使用第一种掩模机制时,单个样本补全效果的可视化差别较小;而使用不规则的掩模机制时,统计学方法更为直接,效果也比较稳定,但补全精度有待提升。深度学习方法在数据整体补全上效果良好,但部分模型如 TCN 和 DCGAN,可能出现模型过拟合的情况,在少量样本上补全效果不稳定。而本文提出的 UMAAtNet 在不同的缺失情况下均有较高的补全精度,效果稳定。

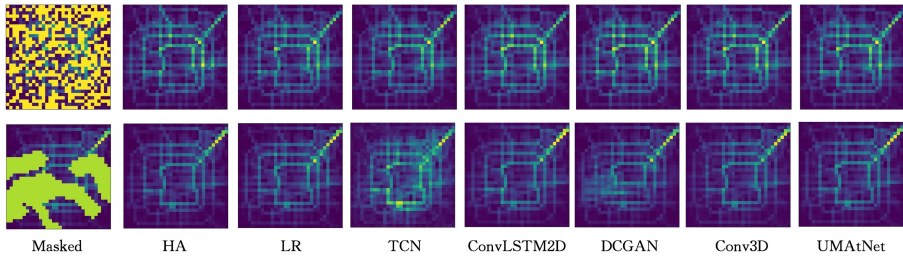


图 10 各模型补全效果的对比图

Fig. 10 Comparison of the completion effects of each model

4.4 不同实验设置的影响

4.4.1 模型组件对实验结果的影响

表 3 中 USAAtNet 表示 U-Net+Self\_Attention,即在初始的 U-Net 结构上加入了 Self\_Attention<sup>[25]</sup>,实现方式为多头注意力机制。由实验结果可以看出,使用该注意力机制能够提高 U-Net 在交通流量补全上的精度,但使用本文改进的 UMaTNet 模型,其建立在本文改进的多视图注意力机制之上,能够从不同的角度得到数据的特征表示,相比前面的方法,补全精度得到较大提升。实验结果如图 11 所示,验证了本文提出的模型组件的有效性。其中,模型细节部分如 UMaTNet+Skip 表示 UMaTNet 加入了跳跃连接的结果,该方法区分了模型的输出,侧重于对每个缺失区域进行补全,提高了数据补全的精度。UMaTNet+Skip+Ext 表示 UMaTNet 同时加入了跳跃连接和外部影响因素的结果,外部因素如天气、节假日等对缺失区域的流量补全提升较小,但从数据影响因素的角度来说,考虑得更加全面。

表 3 模型各组件对实验结果的影响

Table 3 Results with different components of the model (RMSE)

Model/Missing-rate P	10%	20%	30%	40%	50%
U-Net	5.60	6.76	7.54	9.22	12.14
USAAtNet	3.84	5.60	6.20	6.81	7.91
UMaTNet	3.80	5.10	5.98	6.79	7.78
UMaTNet+Skip	3.26	4.56	5.62	6.50	7.53
UMaTNet+Skip+Ext	3.13	4.43	5.44	6.48	7.46

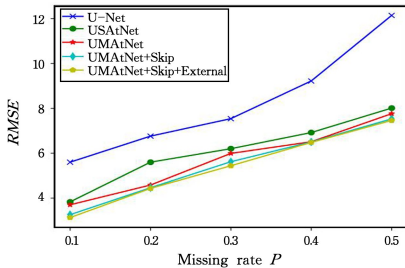


图 11 基于 U-Net 的不同模型组件对补全实验结果的影响

Fig. 11 Data completion experiment comparison results of different model components based on U-Net

4.4.2 不同损失函数的影响

表 4 列出了在 UMaTNet+Ext 模型的基础上,仅修改计算损失函数的组成部分或直接修改损失函数得到的结果。其中 L\_total 表示以式(11)为损失函数的结果,与常规的 MAE 和 MSE 相比,表明以模型的输出分别计算模型对缺失部分的损失和对未缺失部分的损失,效果差于我们直接使用跳跃

连接的方法,即修改计算损失函数的组成部分而非对损失函数进行调整更为直接,也更有效。而在使用跳跃连接之后,采用 MAE 和 MSE 的误差较小,对实验结果的影响也较小。实验对比结果如图 12 所示。

表 4 不同损失函数的结果(RMSE)

Table 4 Results with different Loss function (RMSE)

Loss/Missing-rate P	10%	20%	30%	40%	50%
L_total	3.67	4.78	5.92	6.97	7.92
MAE	3.20	4.45	5.47	6.38	7.46
MSE	3.13	4.43	5.44	6.48	7.46

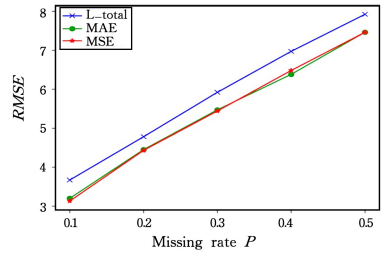


图 12 不同损失函数的实验对比

Fig. 12 Comparison of different loss functions

**结束语** 本文提出了一种时空 U-Net 城市交通流量补全模型,融合了历史时刻的时序数据和空间数据,同时创新性地使用了一种多视图注意力机制,能够从不同的数据视角分析数据特性,利用交通流量的时空特性,对区域的缺失数据进行了有效的补全。在补全模型中融合了多个组件,多视图注意力机制建立在先进的多头注意力方法之上;跳跃连接的方式更有效地针对每个缺失的区域进行补全;外部数据如天气、节假日信息弥补了多模态数据带来的影响。在实验中,我们分析了多视图注意力机制、跳跃连接、外部数据对补全模型效果的影响,实验结果表明本文的补全方法是稳定且有效的。未来,我们希望在图数据结构上利用上述组件的思想,在图数据交通流量预测或补全上进行相应的研究,并进行更多智能交通系统和城市计算的相关应用。

参考文献

[1] LU Y, HU J, XU J, et al. Urban Traffic Flow Forecasting Based on Adaptive Hinging Hyperplanes[C]// International Conference on Artificial Intelligence and Computational Intelligence. Berlin, Heidelberg: Springer, 2009: 658-667.  
 [2] 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.

- [3] SMITH B, SCHERER W, CONKLIN J. Exploring Imputation Techniques for Missing Data in Transportation Management Systems[J]. *Transportation Research Record Journal of the Transportation Research Board*, 2003, 1836: 132-142.
- [4] HAN W G, WANG J F, HU J J. Interpolation Method of Missing Value in Traffic Flow Data[J]. *Journal of Transport Information and Safety*, 2005, 23(1): 39-42.
- [5] BOX G E P, JENKINS G M, REINSELG C, et al. *Time Series Analysis: Forecasting and Control*, 5th Edition[J]. *Journal of the Operational Research Society*, 2015, 22(2): 199-201.
- [6] MING Z, SATISH S, PAWAN L. Matching Patterns for Updating Missing Values of Traffic Counts[J]. *Transportation Planning and Technology*, 2006, 29: 141-156.
- [7] ZHONG M, LINGRAS P, SHARMA S. Estimation of missing traffic counts using factor, genetic, neural, and regression techniques[J]. *Transportation Research Part C Emerging Technologies*, 2004, 12(2): 139-166.
- [8] TAN H, FENG G, FENG J, et al. A Tensor Based Method for Missing Traffic Data Completion[J]. *Transportation Research Part C Emerging Technologies*, 2013, 12(2): 139-166.
- [9] RAN B, TAN H, WU Y, et al. Tensor based missing traffic data completion with spatial-temporal correlation[J]. *Statistical Mechanics and its Applications*, 2016, 446: 54-63.
- [10] TIAN Y, ZHANG K, LI J, et al. LSTM-based traffic flow prediction with missing data[J]. *Neurocomputing*, 2018, 318(27): 297-305.
- [11] KU W C, JAGADEESH G R, PRAKASH A, et al. A Clustering-Based Approach for Data-Driven Imputation of Missing Traffic Data[C]// *IEEE Forum on Integrated and Sustainable Transportation Systems*. IEEE, 2016.
- [12] YU F, WEI D, ZHANG S, et al. 3D CNN-based Accurate Prediction for Large-scale Traffic Flow[C]// *2019 4th International Conference on Intelligent Transportation Engineering (ICITE)*. 2019: 99-103.
- [13] LI Z, ZHENG H, FENG X. 3D Convolutional Generative Adversarial Networks for Missing Traffic Data Completion[C]// *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*. 2018.
- [14] WANG L, LI M, YAN J Q, et al. Urban Traffic Flow Data Recovery Method Based on Generative Adversarial Network[J]. *Journal of Transportation Systems Engineering and Information Technology*. 2018, 18(6): 63-71.
- [15] YUAN Y Y, KANG Y, LI H, et al. Timing traffic flow data completion based on ST-DCGAN[J]. *Computer Engineering and Applications*, 2020, 56(15): 140-146.
- [16] HONG Y L, BIN J, YI X, et al. Coherent Semantic Attention for Image Inpainting [C]// *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019: 4170-4179.
- [17] GUILIN L, FITSUM A, KEVIN J, et al. Image Inpainting for Irregular Holes Using Partial Convolutions [C] // *European Conference on Computer Vision*. 2018(11): 89-105.
- [18] WEN X, LI T, HAN Z, et al. Point Cloud Completion by Skip-attention Network with Hierarchical Folding[C]// *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2020.
- [19] OLAF R, PHILIPP F, THOMAS B. U-Net: Convolutional Networks for Biomedical Image Segmentation [C] // *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, Cham, 2015: 234-241.
- [20] ZHANG J, ZHENG Y, QI D, et al. Predicting Citywide Crowd Flows Using Deep Spatio-Temporal Residual Networks[J]. *Artificial Intelligence*, 2018, 259: 147-166.
- [21] WANG L, GENG X, MA X, et al. Crowd Flow Prediction by Deep Spatio-Temporal Transfer Learning [J]. *arXiv: 1802.00386*, 2018.
- [22] VU D L, VU N T. Vietnamese speech recognition using Dynamic Time Warping and Coefficient of Correlation[C]// *International Conference on Control, Automation and Information Sciences (ICCAIS)*. 2013: 64-67.
- [23] TANG R, KANAMORI R, YAMAMOTO T. Short-Term Urban Link Travel Time Prediction using Dynamic Time Warping with Disaggregate Probe Data[J]. *IEEE Access*, 2019, 99: 1.
- [24] ASHISH V, NOAM S, NIKI P, et al. Attention is All you Need [C]// *Conference and Workshop on Neural Information Processing Systems*. 2017: 5998-6008.
- [25] AHMED K, KESKAR N S, SOCHER R. Weighted transformer network for machine translation[J]. *arXiv: 1711.02132*, 2017.
- [26] ZHU G, ZHANG L, SHEN P, et al. Continuous Gesture Segmentation and Recognition Using 3DCNN and Convolutional LSTM[J]. *IEEE Transactions on Multimedia*, 2018, 21: 1011-1021.
- [27] LEA C, FLYNN M D, VIDAL R, et al. Temporal convolutional networks for action segmentation and detection [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017: 156-165.
- [28] XIONG T, QI Y, ZHANG W B. Short-term Traffic Flow Prediction Based on DCGRU-RF Model for Road Network[J]. *Computer Science*, 2020, 47(5): 84-89.
- [29] RONG B, WU Z H, LIU X H, et al. Flow Prediction of Traffic Stations Based on Spatio-Temporal Multi-Graph Convolutional Network[J]. *Computer Engineering*, 2020, 46(5): 26-33.
- [30] KONG F Y, ZHOU Y F, CHEN G. Traffic Flow Prediction Method Based on Spatio-Temporal Feature Mining[J]. *Computer Science*, 2019, 46(7): 322-326.



**KANG Yan**, born in 1972, Ph.D, associate professor. Her main research interests include software engineering, system optimization, big data processing and mining.