

面向交通流预测的时空编码器-解码器模型

张锦^{1,2}, 皮煜¹, 孙程², 魏叶华^{1*}, 余飞², 姚卫²

(1. 湖南师范大学信息科学与工程学院, 湖南长沙 410081; 2. 长沙理工大学计算机学院, 湖南长沙 410114)

摘要:为了解决许多交通流预测研究方法不能全面地挖掘交通数据中的动态隐藏相关性的问题,研究了动态时空变化特征,提出了一个基于编码器-解码器的交通预测模型。在模型中,编码器和解码器都主要由多头时空注意力机制组成,在两者中间加入了连接注意力机制,以分析路网时空相关性。模型还使用时空嵌入编码与自适应图卷积结合构成的动态嵌入模块来分析节点的动态和静态信息。在两个真实数据集上的实验,证明了该时空模型在长短期流量预测的效果优于其他方法。因此,时空编码器-解码器模型能有效处理复杂的时空序列,提升交通流预测的准确性。

关键词:交通流预测;图卷积神经网络;注意力机制;编码器-解码器

中图分类号:TP391 **文献标志码:**A **文章编号:**1001-2486(2025)03-173-10



Spatial-temporal encoder-decoder model for traffic flow prediction

ZHANG Jin^{1,2}, PI Yu¹, SUN Cheng², WEI Yehua^{1*}, YU Fei², YAO Wei²

(1. College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China;

2. School of Computer Science, Changsha University of Science and Technology, Changsha 410114, China)

Abstract: In order to solve the problem that many traffic flow prediction research methods are unable to comprehensively explore the dynamic hidden correlations in traffic data, the dynamic spatio-temporal variation characteristics were studied and an encoder-decoder-based traffic prediction model was proposed. In the model, both encoder and decoder mainly consisted of multi-head spatio-temporal attention mechanism modules, and a connection attention mechanism was added in between to analyze the spatio-temporal correlations of the road network. The model also used a dynamic embedding module consisting of a combination of both spatio-temporal embedding coding and adaptive graph convolution to analyze the dynamic and static information of nodes. Experiments on two real datasets demonstrate that the spatio-temporal model outperform other models for long- and short-term traffic prediction. Thus, the spatio-temporal encoder-decoder model can effectively handle complex spatio-temporal sequences and improve the traffic flow prediction accuracy.

Keywords: traffic flow prediction; graph convolutional neural network; attention mechanism; encoder-decoder

智慧交通系统(intelligent transportation system, ITS)是智慧城市中的重要组成部分^[1],预测交通流量成为ITS中至关重要的一个环节。交通流的预测包括交通速度、密度、道路利用率、区域流入流出车辆数等特征的预测。通过准确的交通流量预测,路段的拥堵可以被提前警示,以提高交通运输效率。交通流预测的宗旨是交通系统根据历史的交通数据来预测未来的交通状况。近些年来交通预测的研究与应用方面不断扩大,但仍有部分因素导致交通预测成为一个有挑战性的任

务。这些因素包括:

1) 复杂的空间相关性:路网的结构是基于非欧几里得的,不同区域的道路状况也会因地理位置的不同而具有差异性。

2) 复杂的时间相关性:一个传感器附近的交通流量与其历史交通流量息息相关。交通流量会随着时间变化而动态变化。

3) 长时间序列导致的误差传播^[2]:在较长时间序列中预测交通流量时,会随着时间步长增加而导致预测结果与真实情况之间的误差增大。

收稿日期:2023-02-13

基金项目:国家部委基金资助项目(31511010402);湖南省自然科学基金资助项目(2025JJ60410);湖南省综合交通运输大数据智能处理重点实验室开放研究基金资助项目(B202401)

第一作者:张锦(1979—),男,河南信阳人,教授,博士,博士生导师,E-mail:mail_zhangjin@163.com

*通信作者:魏叶华(1979—),男,河南周口人,教授,博士,硕士生导师,E-mail:yehuahn@163.com

引用格式:张锦,皮煜,孙程,等.面向交通流预测的时空编码器-解码器模型[J].国防科技大学学报,2025,47(3):173-182.

Citation:ZHANG J, PI Y, SUN C, et al. Spatial-temporal encoder-decoder model for traffic flow prediction[J]. Journal of National University of Defense Technology, 2025, 47(3): 173-182.

交通流预测方法可以分成基于统计方法、机器学习方法和深度学习的模型。最早的交通流预测模型是统计模型,有历史平均值(historical average, HA)^[3]、自回归综合移动平均值(autoregressive integrated moving average, ARIMA)^[4]、向量自回归(vector autoregressive, VAR)^[5]和支持向量回归(support vector regression, SVR)^[6]等。HA、ARIMA、VAR 为基于统计方法的交通流预测模型,它们使用历史数据进行预测,模型理论较为简单,但是大部分模型都是基于线性基础的,因此在变化大的交通流数据预测中很难体现准确的预测结果。SVR 模型为基于机器学习的预测模型,可以从数据中提取隐藏信息,但模型性能较差,预测精度也不高。

近年来,深度学习模型被广泛用于交通流预测,尤其是基于图结构的深度学习模型。许多科研人员将交通路网建模为一个图结构,并使用图卷积神经网络(graph convolutional network, GCN)^[7]挖掘空间相关性。时间图卷积网络(temporal graph convolutional network, T-GCN)^[8]结合 GCN 和门控递归单元(gated recurrent unit, GRU),其中 GCN 捕获空间相关性,GRU 捕获时间相关性。时空图卷积网络(spatio-temporal graph convolutional networks, STGCN)^[9]通过图对交通网络进行建模,在时间、空间两个层次作用图卷积,用于交通流量预测任务。随后一些模型中被加入注意力机制来建模动态时空特征,基于注意力机制的时空图卷积(attention based spatial-temporal graph convolutional networks, ASTGCN)^[10]提出了基于注意力的时空图卷积网络模型来解决交通流预测问题。然而以上模型都以固定的邻接矩阵来表示交通路网中的道路连通性,因此图波网络(Graph WaveNet)^[11]、自适应图卷积循环网络(adaptive graph convolutional recurrent network, AGCRN)^[12]通过开发一个新的自适应依赖矩阵并通过节点嵌入来学习它,可以捕捉交通历史数据中的隐藏空间依赖。文献[13]利用图结构和编码器-解码器结构,提出了一个基于图扩散注意力的交通流预测模型,反映出了路网传感器之间动态变化的交通流。

时空相关性的捕捉和节点的动态变化特征是时序预测模型中需要关注的重点问题。因此本研究的主要内容是设计一个基于编码器-解码器的交通流预测模型,编码器负责对历史交通流量数据进行编码并输出向量,解码器则将编码器输出的向量转化为用于预测的输出序列。在两个真实

数据集上进行多项实验,目的是完善对数据的时空关系建模,希望对现有的交通流预测模型做出一定的创新和推动。

1 问题描述及研究动机

将交通路网定义为一个有向图结构 $G = (V, E, \mathbf{A})$ 。在图 G 中: V 代表道路网络上的顶点(例如传感器设备),顶点个数 $N = |V|$,构成一个有限集合; E 是代表连接顶点之间具有连通性的一组边; $\mathbf{A} \in \mathbf{R}^{N \times N}$ 为加权邻接矩阵,代表节点之间的连通性和邻近度,是通过道路距离测量得出的结果。

时间步长 t 处的交通状况表示为图形信号 $\mathbf{X}_t$, $\mathbf{X}_t \in \mathbf{R}^{N \times C}$,其中 C 代表交通流量特征的数目(例如,交通量、交通速度等)。

该研究问题可表述为:给定历史时间步长为 P 的历史数据 $\mathbf{X} = (\mathbf{X}_t, \mathbf{X}_{t+1}, \dots, \mathbf{X}_{t+P})$,表示 N 个顶点的观测值, $\mathbf{X} \in \mathbf{R}^{N \times P \times C}$,根据 $\mathbf{X}$ 预测所有顶点的下一个 Q 个时间步长的交通状况 $\mathbf{Y} = (\mathbf{Y}_{t+P}, \mathbf{Y}_{t+P+1}, \dots, \mathbf{Y}_{t+P+Q})$,表示为 $\mathbf{Y} \in \mathbf{R}^{N \times Q \times D}$ 。

考虑大多数基于 GCN 的模型容易忽略交通流数据中同时存在的时间空间相关性以及每个节点唯一的特征^[8-9]。然而编码器-解码器结构和注意力机制相融合模型可以高效处理变长序列,同时专注于每一个节点的时空变化特征。此外,大部分方法使用的是预定义邻接矩阵来代表路网节点之间的连通性,忽略了节点之间的隐藏空间关联性。虽然部分模型^[11-12]通过基于参数学习自适应地捕捉空间相关性,但是它们忽略了节点之间的依赖关系受到时间变化和空间位置变化的影响,因此本模型将时空嵌入编码结合自适应图卷积,将学习到的邻接矩阵随时空变化而自适应改变。

2 研究方法和模型

2.1 时空编码器-解码器模型的总体架构

图 1 是时空编码器-解码器模型的总体架构。输入特征信息通过全连接层增加维度并提高网络的表达能力。编码器和解码器都包含 L 个时空多头注意力机制块,用于提取数据的时空相关性。动态图卷积嵌入模块结合时空嵌入和自适应图卷积,对节点的动态和静态状态分别进行嵌入。连接注意力机制处理短期流量预测模块和编码器的输出,并结合动态图卷积嵌入,将它们转换并送入解码器。所有层的输出维度表示为 D 。下面会详细介绍每个模块。

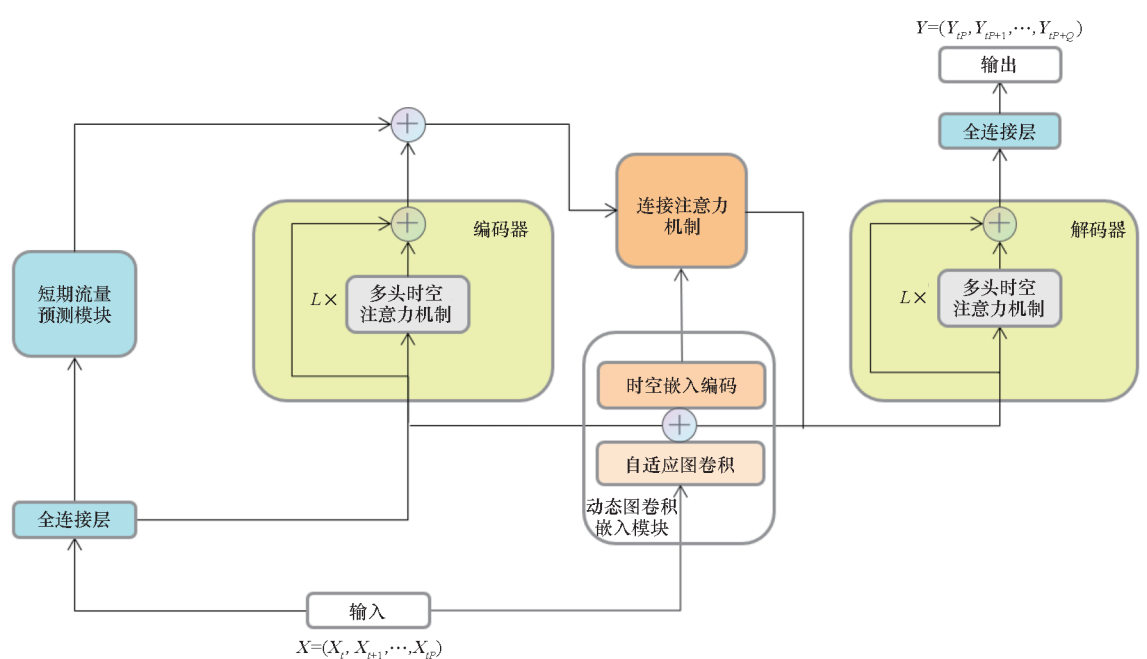


图1 时空编码器-解码器模型的总体架构

Fig. 1 Overall architecture of the spatio-temporal encoder-decoder model

2.2 短期流量预测模块

短期交通流量预测一般指预测未来 30 min 以内的流量。短期预测模块是 ReLU 激活函数^[14]和卷积神经网络 (convolutional neural network, CNN)^[15]构成, 先前的研究已证实 CNN 可有效获取短期时间依赖关系^[16]。本模块设计结构如图 2 所示。把张量 $\mathbf{X} = (\mathbf{X}_t, \mathbf{X}_{t+1}, \dots, \mathbf{X}_{t+p})$ 作为输入, 得到了可初始化预测的一个张量。模块的输出记作 $\mathbf{X}_s \in \mathbf{R}^{N \times P \times D}$, 从输入 $\mathbf{X}$ 到输出 $\mathbf{X}_s$ 的过程记为:

$$\mathbf{X}_s = \text{Short}(\mathbf{X}) + \mathbf{X} \tag{1}$$

式中, Short 代表短期流量输出函数。设置二维卷积核大小为 3×3 , 移动步长为 1, 且卷积层的输入输出通道数都为 64。因为交通流数据是高维度的, 卷积神经网络共享卷积核, 处理高维数据压力较小, ReLU 激活函数能加快网络训练速度, 增加网络非线性, 且残差连接防止了梯度爆炸等现象, 所以基于 CNN 和残差连接的短期流量预测模块能在短期预测中达到很好的效果。

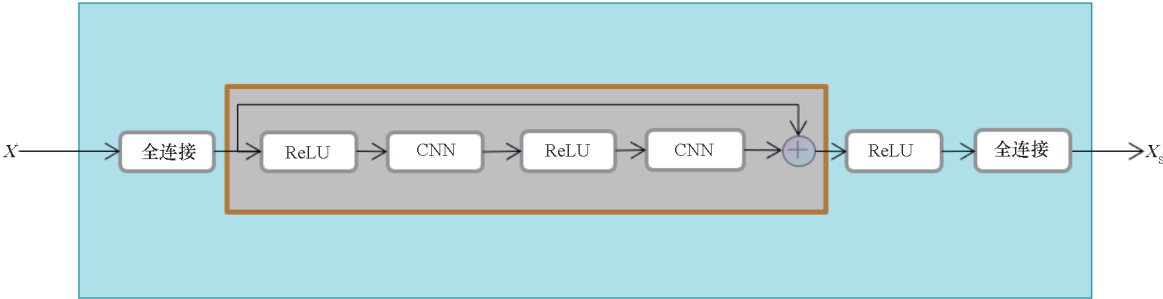


图2 短期流量预测模块结构

Fig. 2 Short-term flow forecasting module structure

2.3 动态图卷积嵌入模块

2.3.1 时空嵌入

受到了 Transformer^[17] 的启发, 并遵循文献[2]设计了时空嵌入编码, 描述路网中的节点在不同时间步长上所代表的位置信息。

空间嵌入将图 G 顶点编码到交通路网图结

构中, 表示为 $\mathbf{e}^s \in \mathbf{R}^{N \times D}$ 。这些向量描述了 N 个传感器顶点的 D 维特征信息, 构成了空间嵌入向量集。这里的空间嵌入表示的是道路的静态情况, 是有预先定义性质的, 所以还需要利用时间步长来构造动态的时间嵌入。

时间嵌入考虑到交通流量的周期性变化, 加

入了 1 d 里的 T 个时间步长和 7 d 里的 7 个时间步长来组成二维时间编码,并使用 One-hot 独热编码来表示为 $\mathbf{e}^t \in \mathbf{R}^{T+7}$,然后通过全连接层映射到维度 D 。

最后,时空嵌入是空间嵌入和时间嵌入相加而成,节点 v_i 在时间步 t_j 的表示为:

$$\mathbf{e}_{v_i, t_j}^{\text{st}} = f_1(\mathbf{e}_{v_i}^s) + f_2(\mathbf{e}_{t_j}^t) \quad (2)$$

式中, f_1, f_2 代表映射维度的全连接层。 N 个节点在 $P + Q$ 个时间步长的情况下表示为 $\mathbf{e}^{\text{st}} \in \mathbf{R}^{(P+Q) \times N \times D}$,它包含了整个输入特征中历史时间以及未来预测的时间片段信息。

2.3.2 自适应图卷积

在本模型中,GCN 直接用于结构化交通图形数据,提取空间特征信息。

为了解决静态的预定义图很难体现相邻路网节点之间真实依赖关系的问题,本设计提出了动态自适应邻接矩阵 $\mathbf{A}_{\text{apt}}$,它能够从数据中自动学习并推断出节点之间隐藏的相互依赖关系。许多研究^[11, 18-19]表明,使用自适应邻接矩阵代替预定义邻接矩阵在交通预测任务中能体现出更多优势。自适应矩阵为:

$$\mathbf{A}_{\text{apt}} = \text{Softmax}(\text{ReLU}(\mathbf{E}_A \mathbf{E}_A^T)) \quad (3)$$

式中, $\mathbf{E}_A \in \mathbf{R}^{N \times D}$ 是随机初始化的可训练参数矩阵,每一行表示节点的嵌入, D 表示嵌入维度,通过 $\mathbf{E}_A$ 与自身的转置相乘,导出了所有节点的空间相关性权重。因某些距离较远的节点之间连接性很弱,所以需要使 ReLU 激活函数来消除弱连接性^[13]。应用 Softmax 函数^[20]对自适应邻接矩阵进行归一化。参考 Xie 等工作^[7],并通过结合自适应邻接矩阵和输入的图结构信息,将 GCN 操作定义为:

$$\mathbf{Z} = \text{ReLU}(\mathbf{A}_{\text{apt}} \mathbf{X} \mathbf{W} + \mathbf{b}) \quad (4)$$

式中, $\mathbf{X} \in \mathbf{R}^{N \times P \times D}$ 是 GCN 含有拓扑信息的输入特征信息, $\mathbf{Z} \in \mathbf{R}^{N \times P \times D}$ 则是 GCN 的输出, $\mathbf{W}$ 是可学习权重矩阵, $\mathbf{b}$ 为偏置项。自适应图卷积可以自动学习交通路网中节点邻域之间的具有变换性的特征信息,提高了计算效率。

最后,通过时空嵌入增强自适应图卷积,提取了交通路网中静态和动态的信息,得到动态图卷积嵌入模块的最终输出 $\mathbf{X}_G \in \mathbf{R}^{N \times Q \times D}$,表示为:

$$\mathbf{X}_G = \mathbf{e}_{v_i, t_j}^{\text{st}} + \mathbf{Z} \quad (5)$$

2.4 编码器-解码器模块

先前的研究图多头注意力网络 (graph multi-attention network, GMAN)^[2]和时空 Transformer 网络 (spatial-temporal Transformer networks, STTN)^[21]等都使用了时空注意力,并将时空嵌入

或位置编码与注意力模块连接。但这些方法还不足以充分挖掘隐藏状态下的时空依赖,因此本模型通过融合动态图卷积嵌入来加强时空注意力机制以解决此问题。

2.4.1 编码器

编码器主要的任务是将历史交通数据转化为一个有隐藏含义的输出向量。2.3 节中的动态图卷积嵌入模块的输出 $\mathbf{X}_G$ 和输入特征 $\mathbf{X}$ 分别送入编码器的多头注意力机制模块中,编码器的输入记为 $\mathbf{X}_{\text{Ein}}$ 。

1) 多头空间注意力机制。交通特征信息融合动态图卷积嵌入能够自适应地捕捉交通路网中传感器之间的相关性,关键思想是在不同的时间步长向不同的顶点动态分配不同的权重。因此,将 $\mathbf{X}_{\text{Ein}}$ 映射到 3 个矩阵: query 矩阵 $\mathbf{Q}_s$, key 矩阵 $\mathbf{K}_s$, value 矩阵 $\mathbf{V}_s$ 。多头注意力机制可以关注多个方面信息,以获取更完整的数据,将 $\mathbf{Q}_s, \mathbf{K}_s, \mathbf{V}_s$ 拆分成维度 $d = D/h$, h 为多头注意力的头数目。空间注意力机制可以描述为:

$$\mathbf{A}_s = \frac{\exp\left(\frac{\mathbf{Q}_s \mathbf{K}_s^T}{\sqrt{D}}\right)}{\sum_{v_i=1}^{V_N} \exp\left(\frac{\mathbf{Q}_s \mathbf{K}_s^T}{\sqrt{D}}\right)} \mathbf{V}_s \quad (6)$$

式中, $\mathbf{A}_s \in \mathbf{R}^{N \times P \times D}$ 为注意力机制得分矩阵;分母 $\sqrt{D}$ 是为了使训练过程中梯度值保持稳定;在交通路网中, V_N 表示传感器 (或节点的数量),通过归一化,模型可以动态地比较不同节点的重要性。下一步,将 h 个头的空间注意力矩阵拼接在一起,得到:

$$\mathbf{MA}_s = \parallel_{i=1}^h (\mathbf{A}_{s1}, \dots, \mathbf{A}_{sh}) \mathbf{W}_0 \quad (7)$$

式中, $\mathbf{W}_0 \in \mathbf{R}^{N \times P \times D}$ 为可训练参数矩阵, $\parallel$ 为拼接操作。

2) 多头时间注意力机制。因为时间相关性会受到交通状态和相应时间段前后的影响,所以主要任务是测量不同时间步长之间的相关性。上述的融合数据信息能够自适应地对不同时间步长之间的线性、非线性相关性进行建模。 T_p 表示时间窗口的长度 (例如过去 P 个时间步),通过归一化,模型可以比较不同时间步的贡献。同理,将时间注意力机制表示为:

$$\mathbf{A}_T = \frac{\exp\left(\frac{\mathbf{Q}_T \mathbf{K}_T^T}{\sqrt{D}}\right)}{\sum_{t_i=1}^{T_p} \exp\left(\frac{\mathbf{Q}_T \mathbf{K}_T^T}{\sqrt{D}}\right)} \mathbf{V}_T \quad (8)$$

$$\mathbf{MA}_T = \parallel_{i=1}^h (\mathbf{A}_{T1}, \dots, \mathbf{A}_{Th}) \mathbf{W}_0 \quad (9)$$

时间注意力机制和空间注意力机制经过门控融

合,得到编码器的多头时空注意力模块输出结果。门控融合过程为:

$$\text{Output}_A = \lambda \odot \mathbf{MA}_S + (1 - \lambda) \odot \mathbf{MA}_T \quad (10)$$

$$\lambda = \sigma(\mathbf{MA}_S \mathbf{W}_1 + \mathbf{MA}_T \mathbf{W}_2 + \mathbf{b}) \quad (11)$$

其中: Output_A 被临时定义为时空注意力机制结合的输出结果; $\odot$ 为两个张量矩阵的点乘运算; $\mathbf{W}_1$, $\mathbf{W}_2$ 为可训练参数矩阵; $\mathbf{b}$ 为偏置项; σ 为 sigmoid 激活函数; λ 为门控单元。门控融合机制动态控制每个顶点和时间步长处的时间和空间相关性。

为了解决梯度消失和可训练参数矩阵的退化问题,编码器中使用残差连接并得到编码器的输出 $\mathbf{X}_{\text{Eout}} \in \mathbf{R}^{N \times P \times D}$ 。为了使解码器能解析更多具有价值的交通数据信息,更精准地预测短期交通流量,将 2.2 节短期流量预测模块的输出 $\mathbf{X}_S$ 与编码器的输出 $\mathbf{X}_{\text{Eout}}$ 相加,得到编码器融合信息序列 $\mathbf{X}_{\text{ES}} \in \mathbf{R}^{N \times P \times D}$ 。

2.4.2 解码器

解码器的任务是预测交通流量,其输入的数据都是 Q 个时间步长。解码器接收从连接注意力机制(下文介绍)中转化的编码器与短期流量预测模块的融合输出,同时合并了动态图卷积嵌入模块的结果,组成了解码器的输入 $\mathbf{X}_{\text{Din}}$ 。在解码器里,时空多头注意力机制的定义与上述编码器中的形式大致相同,唯一改变的是输入的特征信息。动态图卷积嵌入模块的输出 $\mathbf{X}_C$ 与连接注意力机制的输出 $\mathbf{X}_C$ 组成解码器的输入 $\mathbf{X}_{\text{Din}}$,因此 $\mathbf{X}_{\text{Din}}$ 映射的 3 个矩阵变成了:

$$\begin{cases} \mathbf{Q} = \mathbf{X}_{\text{Din}} \mathbf{W}_Q \\ \mathbf{K} = \mathbf{X}_{\text{Din}} \mathbf{W}_K \\ \mathbf{V} = \mathbf{X}_{\text{Din}} \mathbf{W}_V \end{cases} \quad (12)$$

其中, $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V$ 均为可训练参数矩阵。后续工作与编码器的注意力机制同理,得到多头时空注意力机制的输出为 $\mathbf{MA} \in \mathbf{R}^{N \times Q \times D}$ 。数据通过残差连接得到解码器的输出 $\mathbf{X}_{\text{Dout}} \in \mathbf{R}^{N \times Q \times D}$,其经过全连接层并输出接下来 Q 个时间步长的交通流预测数据。

2.5 连接注意力机制模块

连接注意力机制的主要目的是对未来 Q 个时间步长和历史 P 个时间步长之间的直接关系建模,缓解预测中长时间范围内不同预测步长之间的误差传播。2.3.1 节中的时空嵌入模块会被用于连接注意力机制,用于提供路网图形结构和时间信息。

连接注意力机制模块的查询 query、键 key、值 value 分别为 Q 个时间步长的时空嵌入 $\mathbf{e}_Q^{\text{st}} \in \mathbf{R}^{Q \times N \times D}$ 、 P 个历史时间步长的时空嵌入 $\mathbf{e}_P^{\text{st}} \in \mathbf{R}^{P \times N \times D}$ 、编码器和短期流量预测模块的融合输出 $\mathbf{X}_{\text{ES}} \in \mathbf{R}^{N \times P \times D}$ 。因此连接注意力得分矩阵可以描述为:

$$\mathbf{A}_C = \text{Softmax}\left(\frac{(\mathbf{W}_Q \mathbf{e}_Q^{\text{st}})(\mathbf{W}_P \mathbf{e}_P^{\text{st}})^T}{\sqrt{D}}\right) \mathbf{W}_V \mathbf{X}_{\text{ES}} \quad (13)$$

式中, $\mathbf{W}_Q, \mathbf{W}_P, \mathbf{W}_V$ 均为可训练参数矩阵。接下来多头连接注意力机制表示为:

$$\mathbf{MA}_C = \parallel_{i=1}^h (\mathbf{A}_{C1}, \dots, \mathbf{A}_{Ch}) \mathbf{W}_O \quad (14)$$

$\mathbf{MA}_C$ 经过一次全连接层的线性转化后,最终得到了连接注意力机制模块的输出 $\mathbf{X}_C \in \mathbf{R}^{N \times Q \times D}$,将 P 长度的历史数据转化为 Q 长度的预测数据。连接注意力机制的效果如图 3 所示,箭头代表交通图中顶点 V_i 在对应 2 个时间之内的注意力得分。

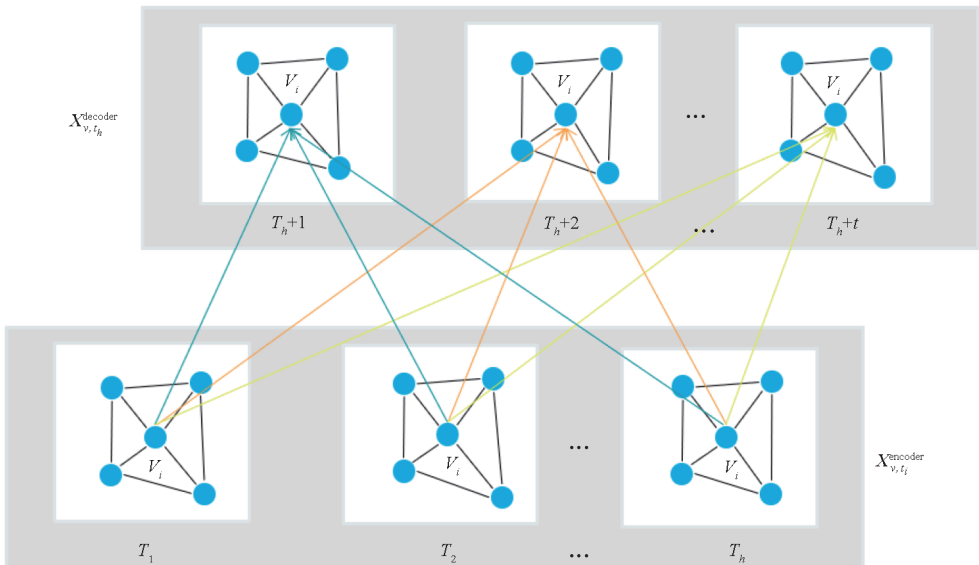


图3 连接注意力机制模块示意图

Fig. 3 Schematic diagram of the connection attention mechanism module

2.6 模型算法训练流程及时间复杂度分析

如算法 1 所示,模型时间复杂度包括:①全连接网络层的时间复杂度(步骤 1、7),大约是 $O(N)$;②短期流量预测模块的时间复杂度(步骤 2),大约是 $O(N)$;③时空嵌入编码的时间成本(步骤 3)大约为 $O((P+Q)ND)$,其中 P 、 Q 、 D 为常量(下文同理),因此时间复杂度约为 $O(N)$;④自适应图卷积模块的时间成本(步骤 3)约为 $O(N^2D^2)$,因此时间复杂度为 $O(N^2)$;⑤多头时空注意力模块的时间复杂度分别约为 $O(N^2DP)$ (步骤 4)和 $O(N^2DQ)$ (步骤 6),因此复杂度约为 $O(N^2)$;⑥连接注意力模块的时间成本约为 $O(PQDN^2)$,所以复杂度约为 $O(N^2)$;⑦损失函数的时间复杂度(步骤 8)大约为 $O(Q)$,所以复杂度记为 $O(1)$ 。因此时空编码器-解码器算法训练过程的总时间复杂度为 $O(N + N + N + N^2 + N^2 + N^2 + 1) \approx O(N^2)$ 。

算法 1 时空编码器-解码器训练过程

Alg. 1 Training process of spatio-temporal encoder-decoder

输入:给定路网图 G 历史时间步长为 P 的历史数据 X
输出:模型的参数

repeat

1. 将历史数据 X 输入全连接网络层得到高维度特征数据 X
2. 通过式(1)处理 X ,得到短期流量预测模块输出 X_s
3. 通过式(2)得到时空嵌入编码 e^{st} ,同时自适应图卷积(式(4))计算输出矩阵 Z ,两者相加得到图卷积嵌入模块的输出 X_c
4. 利用式(6)~(7)和式(8)~(9)分别得到编码器中的多头空间注意力机制和多头时间注意力机制,并通过式(10)~(11)对两者进行门控融合,得到编码器输出 X_{Eout} ,与 X_s 相加得到 X_{ES}
5. e^{st} 和 X_{ES} 接入连接注意力机制模块,通过式(13)~(14)计算出矩阵 X_c
6. 同步步骤 4 中计算注意力机制的方式,将 X_c 和 X_c 送入解码器中的多头时空注意力机制模块,门控融合机制处理两者之后得到解码器输出 X_{Dout}
7. 使用全连接网络层处理解码器输出特征,输出结果为时空编码器-解码器的前向传播输出 Y

8. 以 MAE 定义损失函数 $loss = \frac{1}{Q} \sum_{t=t_{P+1}}^{t_{P+Q}} |Y_t - \hat{Y}_t|$,

Adam 优化器^[22]反向传播优化模型参数

until 模型收敛

return 优化后的模型参数

3 实验与分析

3.1 数据集介绍

为了评估时空编码器-解码器模型的性能,本模型在两个不同的数据集上进行了实验:①PEMS-BAY 数据集^[22],它包含由加州交通局性能测量系统(performance measurement system, PEMS)从海湾地区收集的 2017 年 1 月 1 日至 2017 年 5 月 31 日期间 325 个传感器的 5 个月交通信息。②METR-LA 数据集^[22],它包含从 2012 年 3 月 1 日至 2012 年 6 月 30 日在 207 个传感器上记录的 4 个月的交通速度数据,这些数据是从洛杉矶县高速公路的环形探测器收集的。

3.2 数据处理及实验设置

采用了与 GMAN^[2]中相同的数据预处理程序。在两个数据集中,时间步长 1 表示 5 min,则 1 d 的时间步长为 288。使用 ZScore 方法对数据进行归一化,训练集:验证集:测试集 = 7 : 1 : 2。为了构建道路网络图,将每个交通传感器视为一个顶点,计算传感器之间成对的道路网络距离来构造邻接矩阵:

$$A_{ij} = \begin{cases} \exp\left(-\frac{\text{dist}(v_i, v_j)^2}{\delta^2}\right), & \exp\left(-\frac{\text{dist}(v_i, v_j)^2}{\delta^2}\right) \geq e \\ 0, & \text{其他} \end{cases} \quad (15)$$

其中, A_{ij} 表示传感器 v_i 和 v_j 之间的权重, $\text{dist}(v_i, v_j)$ 为传感器 v_i 和 v_j 之间的道路网络距离, δ 为距离的标准偏差, e 为阈值(设置为 0.1)。

实验在 GPU 环境下进行,用 Python 3.7 和 Pytorch 1.12.1 实现,选择的 GPU 是 Tesla T4。遵循先前的工作,使用 $P = 12$ 的历史步长来预测下一个 $Q = 3, 6, 12$ (15 min, 30 min, 60 min) 的交通状况。使用 Adam 优化器^[23]和 0.001 的初始学习率训练模型, batchsize 选择 32。所有输出层维度 $D = 64$ 。多头注意力机制中,设置头部 $h = 8$,每个头的维度为 8 ($d = D/h$)。编码器和解码器中多头时空注意力块的数目都设置为 3。采用平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)^[22]以及平均绝对百分比误差(mean absolute percentage error, MAPE)三种广泛使用的指标来评估预测性能,它们的定义如下,其中 Y_t 是实际值,而 $\hat{Y}_t$ 是预测值。

1) MAE 是预测值和实际值之间绝对差值的平均值:

$$MAE = \frac{\sum_{t=1}^Q |Y_t - \hat{Y}_t|}{Q} \tag{16}$$

2)RMSE 是预测值和实际值之间的差异的度量:

$$RMSE = \sqrt{\frac{\sum_{t=1}^Q (Y_t - \hat{Y}_t)^2}{Q}} \tag{17}$$

3)MAPE 是绝对百分比误差的平均值:

$$MAPE = \frac{1}{Q} \sum_{t=1}^Q \frac{|Y_t - \hat{Y}_t|}{Y_t} \tag{18}$$

3.3 模型对照基线

将时空编码器-解码器模型与以下模型进行性能比较。ARIMA^[4]是使用时间序列数据之间的相关性来预测交通流量的经典统计方法模型;SVR^[6]是用支持向量机来进行交通流量预测的机器学习方法模型;扩散卷积递归神经网络模型(diffusion convolutional recurrent neural network model, DCRNN)^[24]是扩散卷积神经网络,融合时

间空间依赖性来预测交通流量;STGCN^[9]是使用空间域的图卷积和时间域的一维卷积来进行流量预测;ASTGCN^[10]是使用时空注意力机制与图卷积融合的交通流量预测模型;Graph WaveNet^[11]在图卷积中引入自适应矩阵,能动态学习交通数据的空间依赖关系;图多注意力网络模型(graph multi-attention network model, GMAN)^[2]使用编码器-解码器,结合注意力机制以及时空嵌入编码预测交通流量。

3.4 实验结果

在 METR-LA 和 PEMS-BAY 数据集对比了时空编码器-解码器模型与上述对照基线分别在时间步为 3(15 min)、6(30 min)、12(60 min)的性能,结果见表 1、表 2。由表 1、表 2 的结果得知:①深度学习模型的性能优于统计方法和机器学习方法的模型,证明了深度学习神经网络的强大以及其在线性、非线性交通数据建模方面的能力;②图卷积网络的模型(包括 DCRNN、STGCN、ASTGCN、Graph WaveNet、时空编码器-解码器)

表 1 时空编码器-解码器在 METR-LA 数据集与对照基线的性能对比结果

Tab.1 Performance comparison results on METR-LA dataset of spatial-temporal encoder-decoder model and baselines

模型	15 min			30 min			60 min		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
ARIMA	3.99	8.21	9.60	5.15	10.45	12.70	6.9	13.23	17.40
SVR	3.99	8.45	9.30	5.05	10.87	12.10	6.72	13.76	16.70
DCRNN	2.77	5.38	7.30	3.15	6.45	8.80	3.60	7.60	10.50
STGCN	2.88	5.74	7.62	3.47	7.24	9.57	4.59	9.40	12.70
ASTGCN	2.75	5.62	7.51	3.31	6.98	9.32	4.52	9.24	12.62
Graph WaveNet	2.69	5.15	6.90	3.07	6.22	8.37	3.53	7.37	10.01
GMAN	2.69	5.55	7.42	3.15	6.78	9.02	4.03	8.11	11.72
时空编码器-解码器	2.58	4.89	6.68	2.85	5.88	8.13	3.26	7.02	9.45

表 2 时空编码器-解码器在 PEMS-BAY 数据集与对照基线的性能对比结果

Tab.2 Performance comparison results on PEMS-BAY dataset of spatial-temporal encoder-decoder model and baselines

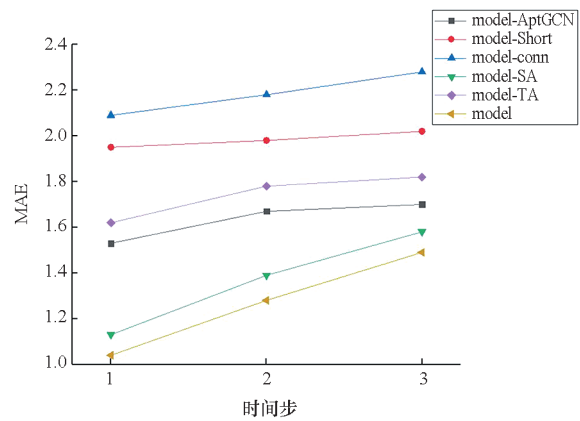
模型	15 min			30 min			60 min		
	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
ARIMA	1.62	3.30	3.50	2.33	4.76	5.40	3.38	6.50	8.30
SVR	1.85	3.59	3.80	2.48	5.18	5.50	3.28	7.08	8.12
DCRNN	1.38	2.95	2.90	1.74	3.97	3.90	2.07	4.74	4.90
STGCN	1.36	2.96	2.90	1.81	4.27	4.17	2.49	5.69	5.79
ASTGCN	1.32	2.78	2.75	1.75	3.98	3.95	2.32	5.41	5.51
Graph WaveNet	1.30	2.74	2.73	1.63	3.70	3.67	1.95	4.52	4.63
GMAN	1.34	2.82	2.81	1.62	3.72	3.63	1.86	4.32	4.31
时空编码器-解码器	1.27	2.43	2.67	1.56	3.22	3.36	1.78	3.95	3.89

的性能优势,表明了交通路网的图结构信息对交通预测十分重要;③ASTGCN 和 GMAN 都使用了时空注意力机制,且相对 DCRNN、STGCN 性能较好,证明了注意力机制捕获时空相关性的重要性;④相比较于基线模型,本模型体现了最优的性能,且在长期流量预测和短期流量预测中都体现出优异的效果。因此时空编码器-解码器模型更适合分析并预测交通流量数据。

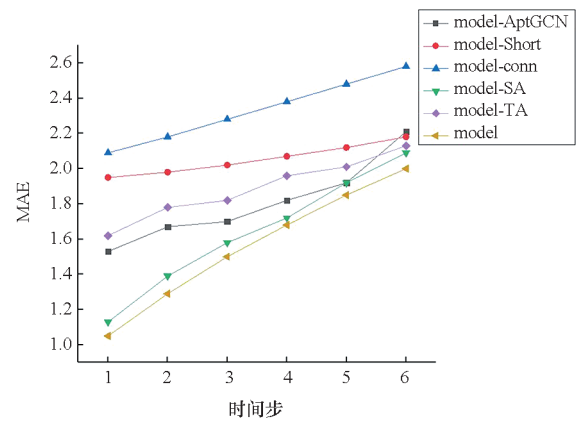
3.5 消融实验

为了验证时空编码器-解码器中每个模块的重要性,在 PEMS-BAY、METR-LA 数据集上将模型中的各个模块分别移除再进行性能测试实验。分别移除自适应图卷积模块、短期流量预测模块、连接注意力模块、空间注意力模块、时间注意力模块进行实验,并依次命名为 model-AptGCN、model-Short、model-conn、model-SA、model-TA、model,本模型记

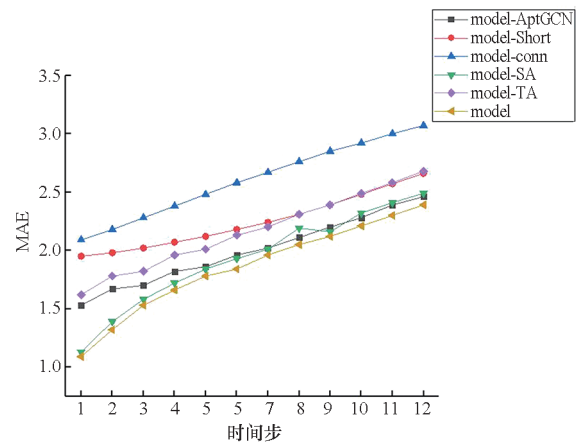
为 model。使用 MAE 来测评。实验结果如图 4 所示,图 4 表明:①时空编码器-解码器每个时间步长的性能始终优于上述 5 个已拆分的模型,这表明时空注意力机制结合自适应图卷积能在建模较复杂的时空相关性方面展现出出色的性能;②当时空编码器-解码器移除短期流量预测模块时,可以见到其预测 30 min 内的流量精度有明显下降,证明了短期流量预测模块的有效性;③当时空编码器-解码器模型移除连接注意力机制模块时,其性能在所有对比实验中表现最差,这也体现了连接注意力机制能非常有效地缓解长时间范围内的误差传播的影响,大幅提高流量预测结果的精确度;④移除空间注意力模块,相比其他移除的模块对预测精度的影响最小,由此可知,在自适应图卷积捕获了空间信息之后,用空间注意力能起到更好的效果,两者结合能更全面地捕获时空相关性。



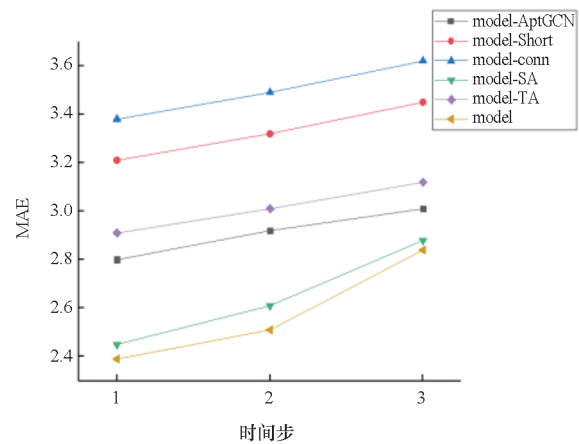
(a) PEMS-BAY 15 min 的 MAE
(a) 15 min MAE on PEMS-BAY



(b) PEMS-BAY 30 min 的 MAE
(b) 30 min MAE on PEMS-BAY



(c) PEMS-BAY 60 min 的 MAE
(c) 60 min MAE on PEMS-BAY



(d) METR-LA 15 min 的 MAE
(d) 15 min MAE on METR-LA

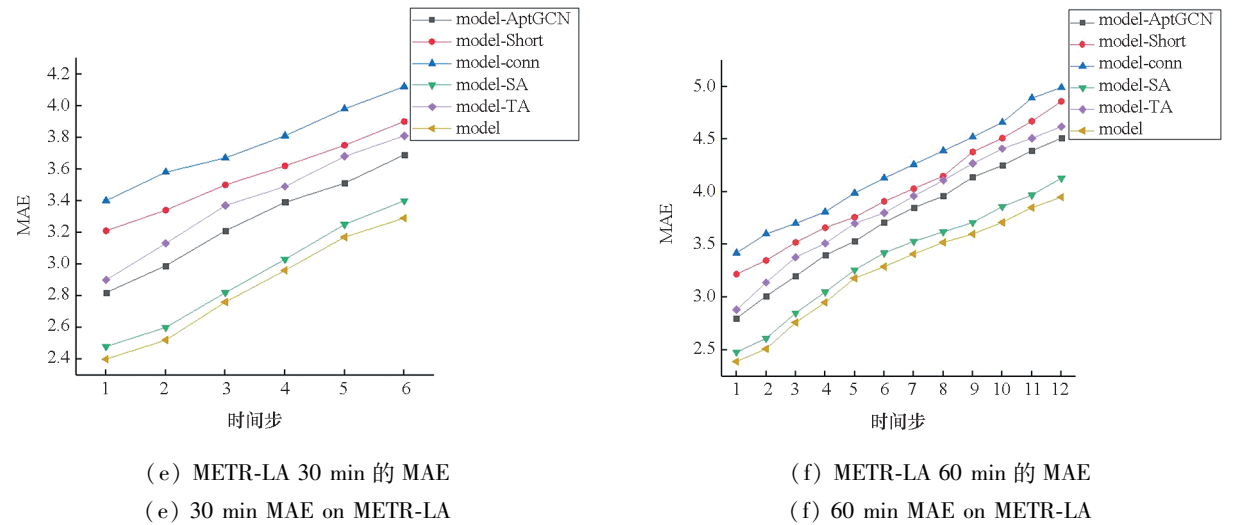


图 4 消融实验结果

Fig.4 Results of ablation experiments

3.6 时空多头注意力机制的特征融合方式评估

为了研究时间注意力、空间注意力机制的特征融合方式对模型性能的影响,本实验分别使用拼接、相加、门控融合机制三种方式评估了它们对模型预测性能的影响。表 3 展示了时空编码器-解码器模型中时空注意力不同的组合

方式预测未来不同时间段流量的性能对比结果,从中可知,使用门控融合机制在两个数据集中表现都为最好,因此门控融合机制能将时空注意力处理的数据信息充分结合,相比与其他两种方式能最大限度地保证信息传输的完整性。

表 3 时空注意力不同组合方式的预测性能对比

Tab.3 Comparison of the prediction performance different combinations of spatio-temporal attention

数据集	融合方式	15 min			30 min			60 min		
		MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
METR-LA	拼接	2.73	5.25	7.12	3.13	6.12	8.62	3.47	7.66	10.01
	相加	2.62	5.16	6.95	3.07	6.06	8.26	3.31	7.28	9.78
	门控融合机制	2.58	4.89	6.68	2.85	5.88	8.13	3.26	7.02	9.45
PEMS-BAY	拼接	1.30	2.71	2.75	1.63	3.66	3.67	1.89	4.01	4.02
	相加	1.29	2.69	2.78	1.59	3.48	3.63	1.81	3.96	3.97
	门控融合机制	1.27	2.43	2.67	1.56	3.22	3.36	1.78	3.95	3.89

4 结论

本研究提出了一种面向交通流预测的时空编码器-解码器模型,使用图结构和多头时空注意力机制,全面地对交通路网进行时空建模。结合不同实验,在两个真实世界的数据集上进行实验评估之后,发现该模型不仅能很好地同时提取复杂时空相关性,而且动态图卷积嵌入模块可以分析路网图中的动态和静态数据信息,提高计算效率并节省计算资源。实验结果表明,时空编码器-解码器模型的性能比 7 个基线模型更优越。

参考文献 (References)

[1] YE J X, ZHAO J J, YE K J, et al. How to build a graph-based deep learning architecture in traffic domain: a survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(5): 3904–3924.

[2] ZHENG C P, FAN X L, WANG C, et al. GMAN: a graph multi-attention network for traffic prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 1234–1241.

[3] 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 J, GUAN W. A summary of traffic flow forecasting methods[J]. Journal of Highway and Transportation Research

- and Development, 2004, 21(3): 82–85. (in Chinese)
- [4] WILLIAMS B M, HOEL L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results [J]. Journal of Transportation Engineering, 2003, 129(6): 664–672.
- [5] CHEN C, PETTY K, SKABARDONIS A, et al. Freeway performance measurement system: mining loop detector data[J]. Transportation Research Record: Journal of the Transportation Research Board, 2001, 1748(1): 96–102.
- [6] WU C H, HO J M, LEE D T. Travel-time prediction with support vector regression [J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4): 276–281.
- [7] XIE Y C, XU Z, ZHANG J T, et al. Self-supervised learning of graph neural networks: a unified review [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(2): 2412–2429.
- [8] ZHAO L, SONG Y J, ZHANG C, et al. T-GCN: a temporal graph convolutional network for traffic prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(9): 3848–3858.
- [9] YU B, YIN H T, ZHU Z X. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting [C]//Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, 2018: 3634–3640.
- [10] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922–929.
- [11] WU Z H, PAN S R, LONG G D, et al. Graph WaveNet for deep spatial-temporal graph modeling [C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, 2019: 1907–1913.
- [12] BAI L, YAO L N, LI C, et al. Adaptive graph convolutional recurrent network for traffic forecasting [C]//Proceedings of the 34th International Conference on Neural Information Processing Systems, 2020: 17804–17815.
- [13] SU J, JIN Z F, REN J, et al. GDFormer: a graph diffusing attention based approach for traffic flow prediction [J]. Pattern Recognition Letters, 2022, 156: 126–132.
- [14] ZHOU X, ZHANG W Z, XU H, et al. Effective sparsification of neural networks with global sparsity constraint [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 3598–3607.
- [15] LI Z W, LIU F, YANG W J, et al. A survey of convolutional neural networks: analysis, applications, and prospects [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(12): 6999–7019.
- [16] ZHANG J B, ZHENG Y, QI D K. Deep spatio-temporal residual networks for citywide crowd flows prediction [J]. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 2017, 31(1): 1655–1661.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6000–6010.
- [18] CAO S Q, WU L B, WU J, et al. A spatio-temporal sequence-to-sequence network for traffic flow prediction [J]. Information Sciences, 2022, 610: 185–203.
- [19] ZHENG C P, FAN X L, PAN S R, et al. Spatio-temporal joint graph convolutional networks for traffic forecasting [J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(1): 372–385.
- [20] GOODMAN J. Classes for fast maximum entropy training [C]//Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001: 561–564.
- [21] ZHOU H, REN D C, XIA H X, et al. AST-GNN: an attention-based spatio-temporal graph neural network for interaction-aware pedestrian trajectory prediction [J]. Neurocomputing, 2021, 445: 298–308.
- [22] 童凯南, 林友芳, 刘军, 等. 面向时空交通栅格流量预测的3D通道注意力网络 [J]. 国防科技大学学报, 2022, 44(3): 41–49.
- TONG K N, LIN Y F, LIU J, et al. 3D channel-wise attention network for spatio-temporal traffic raster flow prediction [J]. Journal of National University of Defense Technology, 2022, 44(3): 41–49. (in Chinese)
- [23] KINGMA D P, BA J L. Adam: a method for stochastic optimization [C]//Proceedings of the 3rd International Conference for Learning Representations, 2015.
- [24] LI Y G, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting [C]//Proceedings of the International Conference on Learning Representations, 2018.