

# 面向动态高速路网的时空注意力网络交通速度预测

邹国建<sup>1, 2</sup>, 赖子良<sup>1, 2</sup>, 李晔<sup>1, 2</sup>▲

(同济大学 同济大学道路与交通工程教育部重点实验室, 上海 201804; 同济大学 交通运输工程学院, 上海 201804)

**摘要:** 高速公路是人们日常出行和货物运输的重要途径之一, 交通速度是影响高速路网通行效率和安全的重要指标。精准预测高速路网交通速度可以减少交通事故和通行时间, 提前为交通控制提供有价值的参考信息, 对高速公路管理具有重要意义。本文提出了一种由数据和长期预测任务驱动的面向动态高速路网的时空注意力网络交通速度预测模型(ST-ANet)。ST-ANet 由三部分组成, 基于空间注意力机制的图注意力网络(GAN)、基于多头自注意力机制(Multi-head self-attention)的时间注意力网络和长短期记忆网络(LSTM)。此外, 本文还使用一些提升模型表现的技巧, 包括密集连接和层归一化方法。首先, 使用 GAN 提取高速路网的动态空间关联特征; 其次, 使用 LSTM 提取输入数据的时间关联特征; 最后, 使用时间注意力机制关注历史输入数据和预测值之间的相关性。实验验证采用中国宁夏省银川市高速路网监测数据。对比最新的 GCN-LSTM 模型, ST-ANet 模型对于未来 1 个小时、2 个小时和 3 个小时内高速路网交通速度的预测误差 MAE 分别降低了 4.0%、3.6%和 3.9%。

**关键词:** 交通工程; 高速路网交通速度预测; ST-ANet 预测模型; 时间注意力; 空间注意力; 长短期记忆网络

开放科学(资源服务)标志码(OSID):



## ST-ANet: Speed prediction of spatio-temporal attention network for dynamic highway network

ZOU Guojian<sup>1,2</sup>, LAI Ziliang<sup>1,2</sup>, LI Ye<sup>1,2</sup>▲

(1. The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China; 2. College of Transportation Engineering, Tongji University, Shanghai 201804, China)

**Abstract:** The highway is one of the crucial ways for people's daily travel and goods transportation, and traffic speed is an important indicator that affects the efficiency and safety of highway network traffic. Accurate prediction of highway network traffic speed can reduce traffic accidents and transit time, provide valuable reference information for traffic control in advance, and is of great significance to highway management. This paper proposes a speed prediction of spatio-temporal attention network (ST-ANet) for dynamic highway network driven by data and long-term prediction tasks. ST-ANet consists of three parts, the graph attention network (GAN) based on the spatial attention mechanism, the temporal attention network based on the multi-head self-attention mechanism, and the long short-term memory network (LSTM). In addition, this paper also uses some tricks to improve model performance, including dense connections and layer batch normalization methods. First, use GAN to extract the dynamic spatial correlation features of the highway network; second, use LSTM to extract the temporal correlation features of the highway network; finally, use the temporal attention mechanism to focus on the correlation between historical input data and predicted values. The evaluation experiments use the monitoring data of the highway network in Yinchuan City, Ningxia Province, China. Compared with the latest GCN-LSTM model, the ST-ANet model reduces the prediction error MAE of the highway network traffic speed in the next 1 hour, 2 hours, and 3 hours by 4.0%, 3.6%, and 3.9%, respectively.

**Key words:** transportation engineering; highway network traffic speed prediction; ST-ANet prediction model; temporal attention; spatial attention; long short-term memory network

**DOI:** 10.19678/j.issn.1000-3428.0063777

基金项目: 国家自然科学基金, 为明天城市的清洁空气融资: 通过土地增值回馈确保城市可持续发展、提高城市空气质量的潜力, (71961137006); 自然科学基金, 智慧公路云平台路网管控关键技术研究联合攻关, (kh0160020213294)

第一作者简介: 邹国建(1993-), 同济大学博士研究生. 研究方向: 智慧高速运行云平台关键算法. E-mail: 2010768@tongji.edu.cn.

▲通信作者: 李晔(1974-), 同济大学教授, 博士. 研究方向: 智能交通、交通能源与环境. E-mail: JamesLI@tongji.edu.cn.

## 0. 引言

高速公路速度预测对于智能交通系统(ITS)至关重要,可为出行者和交通管理部门提前提供有用的交通信息<sup>[1]</sup>。目前高速公路交通速度预测研究主要聚焦于短时预测,对于长时间速度预测能力略显不足<sup>[2]</sup>。根据目前交通速度预测研究的特点,可以将现有的研究方法分为三类:统计方法、传统机器学习方法和深度学习方法。

统计方法成功应用于交通速度预测任务,包括历史平均模型(HA)和自回归综合移动平均模型(ARIMA)<sup>[3-5]</sup>。HA使用同一时间的历史数据的平均值作为未来预测任务中相同时刻的预测值<sup>[3]</sup>。ARIMA作为传统的时间序列预测方法,它结合了移动平均和自回归分量来对历史时间序列数据进行建模<sup>[4-5]</sup>。然而,由于交通速度具有非线性的性质,参数化方法基于先验知识、理论假设和简单的数学统计,在精准预测交通速度方面表现不足。

传统机器学习方法缓解了统计方法遇到的难题,它们提取交通大数据中的非线性特征来提高预测的精准度<sup>[6-10]</sup>。如, Vanajakshi 等提出了一种支持向量机(SVM)的回归技术,用于交通速度的短期预测<sup>[6]</sup>。Jiang 等使用隐马尔可夫模型(HMM)来表现单个车辆的速度与路段交通速度之间的统计关系<sup>[7]</sup>。Shin 等提出了一种基于随机模型的车速预测算法,该算法使用具有速度约束的马尔可夫链作为基础<sup>[8]</sup>。Zhang 等提出了一种基于高阶多元马尔可夫模型的交通因素状态网络模型(TFSN)来建立速度与相关因素之间的关系<sup>[9]</sup>。传统机器学习方法主要提取浅层数据特征,无法对复杂的交通路网时空数据特征进行深度建模<sup>[2, 7-9]</sup>。

近年来,相关研究证明,相比传统机器学习算法,深度学习技术在处理复杂的非线性交通数据优势更加明显<sup>[11-14]</sup>。Csikós 等使用人工神经网络(ANN)进行交通速度预测<sup>[11]</sup>。Jia 等提出了深度信念网络(DBN)模型用于短时交通速度预测<sup>[12]</sup>。Tang 等提出了一种基于改进模糊神经网络(FNN)的交通速度预测模型<sup>[13]</sup>。然而,这些方法处理复杂的交通路网时空数据特征能力有限,因此迫切需要更为复杂的深度学习方法。

循环神经网络(RNNs)是一种用于时间序列预测任务的深度学习方法,可以有效提取数据的时间关联特征<sup>[15]</sup>。目前,许多交通速度预测任务使用 RNNs 作为时序特征提取器来提高预测精准度<sup>[16-20]</sup>。如,Gu 等建立了一种基于 LSTM 和 GRU 的新型融合深度学习(FDL)模型,以捕捉车道短时速度预测的时空特征<sup>[17]</sup>。Wang 等使用双向长短期记忆神经网络(Bi-LSTM NN)对每个关键路段进行建模,使用堆叠 Bi-LSTM 层来合并时间信息<sup>[18]</sup>。上述方法使用 RNN 作为特征提取器提取交通数据的时间相关性,但是忽略数据的空间相关性对预测产生的影响。为了解决 RNNs 遇到的问题,基于 CNNs 的时空预测模型被广泛应用于交通速度预测任务中<sup>[21-27]</sup>。如,为了捕捉复杂的动态交通信息,Zhou 等提出了一种称为时空深度张量神经网络(ST-DTNN)的速度预测方法,主要用于混合道路类型的大规模城市网络<sup>[25]</sup>。Yang 等提出基于路径的速度预测神经网络(PSPNN)由 CNN 和双向 LSTM(Bi-LSTM)网络组成,提取历史数据的时空特征,实现基于路径的速度预测<sup>[26]</sup>。Zang 等提出了一种基于 CovLSTM 的多尺度时空特征学习网络(MSTFLN),用于高架公路长期交通速度预测的研究任务<sup>[27]</sup>。然而,传统的 CNNs 只适用于欧几里德空间结构数据,而交通数据具有非几里德空间性质。因此,基于 CNNs 的交通速度数据空间特征提取存在较大的缺陷。

最新研究将 CNNs 扩展到可以处理非欧几里德空间结构数据的图卷积神经网络(GCNs),并且已成功应用于交通预测任务中,包括交通速度预测<sup>[28-31]</sup>。高速路网中每条道路之间的相关性随着时间动态变化,而 GCNs 受限于捕获高速路网的动态空间相关性。图注意力网络(GANs)优势在于解决动态的空间相关性问题,这给本文研究带来了新的灵感<sup>[32]</sup>。本文提出了一种基于时空注意力机制的动态高速路网速度预测模

型（ST-ANet），并且设计了一个基于 GAN 和 LSTM 的高速公路网络时空特征提取块（GLSTM-block）。首先，使用基于空间注意力机制的 GAN 网络提取高速路网的动态空间相关性；其次，使用 LSTM 专门提取输入数据和 GAN 输出特征的时间相关性；最后，使用时间注意力机制计算历史输入数据和预测值之间的时间相关性。实验结果表明，ST-ANet 的表现超越了所有的基线模型并达到了最优。

## 1. 高速公路速度数据

本文的研究区域为中国宁夏回族自治区银川市滨湖新区的高速路网，如图 1 所示，共采集 8 个高速公路收费站的 ETC 行车数据，49 条高速公路路段的交通速度数据。高速公路交通速度数据包括速度、时间和道路位置三个因素，时间跨度为 2020 年 5 月 1 日至 2020 年 8 月 31 日。每隔一小时采样一次，得到交通速度数据的时间序列形式  $\{x_{t_i} | -\infty < t_i < +\infty\}$ ，其中  $x_{t_i} \in R^{N \times 3}$ ， $N$  表示道路的条数。道路位置不随时间变化，共有 49 条道路，即 49 个位置索引。在实验中，本文使用 80% 的数据作为训练集，20% 的数据作为测试集。

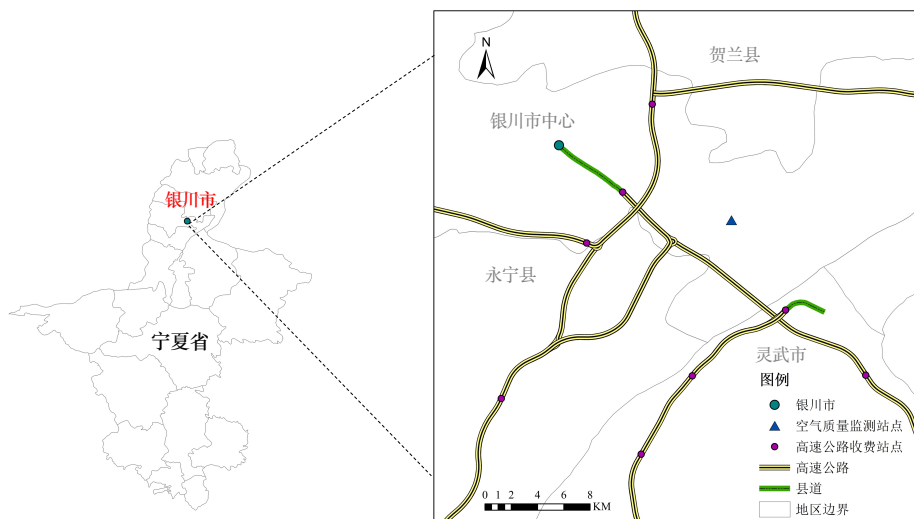


图 1 研究区域  
Fig. 1 Study area

## 2. 面向动态高速路网的时空注意力网络速度预测

### 2.1 问题定义

本文的研究目标是，预测高速公路网络中每一条道路未来一段时间内的交通速度情况。将输入的高速路网图定义为  $G = (V, E, A)$ ，其中  $V$  代表节点集， $E$  表示边集， $A \in R^{N \times N}$  表示邻接矩阵， $N$  表示节点的数量。需要特别强调的是高速路网图，本文将每条道路抽象为图中的节点，将道路的连接抽象为边。如图 2 (a) 所示，每条线段代表一条高速公路，每条道路都可以映射到图 2 (b) 中的图网络节点。邻接矩阵  $A$  表示道路之间连接关系，1 表示两条道路之间存在连接，反之亦然，0 表示没有连接。假设输入时间步长为  $P$ ，预测时间步长为  $Q$ ，且  $t_i \in \{t_1, \dots, t_P, \dots, t_{P+Q}\}$ 。本文研究问题的核心是如何揭示高速路网交通数据的时空相关性。因此，本文使用 ST-ANet 来学习隐藏在公路交通数据中的时空特征，实现高速路网交通速度的精准预测。

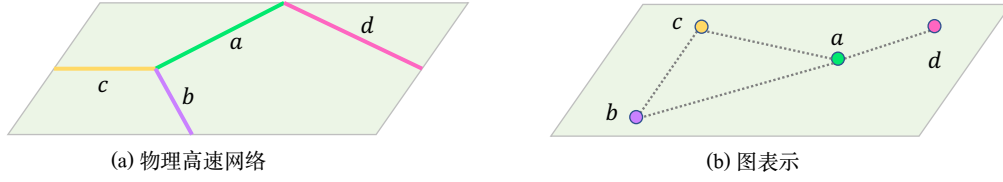


图 2 (a) 有四条道路的高速公路网络示例。每条道路由不同的字符和颜色代表。(b) 高速公路网络用图表示，道路用节点表示

Fig. 2 (a) Example of highway network with four roads. Each road is represented by a different characters and colors. (b) The highway network can be represented as a graph, and roads can be represented as nodes

## 2.2 模型框架

针对时空特征提取和长期公路交通速度预测，本文提出了 ST-ANet 预测模型，如图 3 所示。实现该方法需要两个步骤：

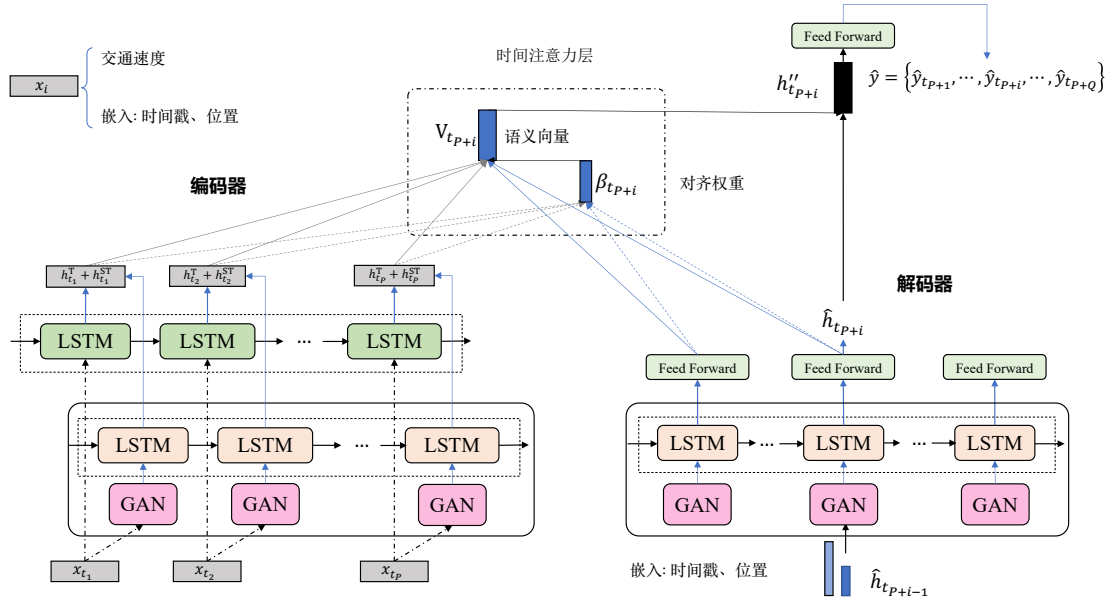


图 3 ST-ANet 预测模型框架。编码器提取输入数据的时空特征，解码器预测未来高速公路交通速度

Fig. 3 The framework of the proposed ST-ANet. The encoder extracts the spatiotemporal features of the input data, and the decoder predicts the future highway traffic speed

**编码器** 公路交通数据  $x = \{x_{t_1}, \dots, x_{t_p}\}$ ,  $x_{t_i} \in R^{N \times d_{model}}$  以流水方式馈入到 ST-ANet 中，通过 GLSTM-Block 学习每个时间步的时空特征。GAN 提取输入交通数据的动态空间相关性，使用 LSTM 提取 GAN 输出的空间特征的时间相关性。

**解码器** 基于编码器输出，使用基于 GLSTM-Block 的解码器来预测长期高速公路交通速度  $\hat{y} = \{\hat{y}_{t_{p+1}}, \dots, \hat{y}_{t_{p+Q}}\}$ ,  $\hat{y}_{t_{p+i}} \in R^{N \times 1}$ 。在整个解码器预测过程中，使用时间注意力机制来计算历史输入数据与每个时间步预测值之间的相关性。

对于本文提出的 ST-ANet，为整个网络添加了密集连接和层归一化技巧，以防止网络信息丢失和内部协变量偏移问题。在以下章节中，将详细描述 ST-ANet 模型每个部分的实现过程。

## 2.3 嵌入层

每个节点包含三类数据信息，包括位置信息、时间戳信息和交通速度信息。输入变量  $x_{t_i} = (e_p^{t_i}, e_s^{t_i}, e_t^{t_i})$  可拆分为三种类型的嵌入表示，包括位置嵌入  $e_p^{t_i} \in R^{N \times d}$ 、交通

速度嵌入  $e_s^{t_i} \in R^{N \times d}$  和时间戳嵌入（小时、周、月）  $e_t^{t_i} \in R^{N \times d}$ 。时间戳和位置嵌入方法与 Transformer 的嵌入方法相同，通过 one-hot 映射到稠密矩阵<sup>[33]</sup>。交通速度嵌入方法通过线性变换完成。对于每一个类型数据嵌入，数据维度可表示为  $d = d_{model}/3 = 64$ 。

## 2.4 GLSTM-Block

如图 4 所示，GLSTM-Block 由两个关键部分组成，动态空间特征提取器和时间特征提取器。

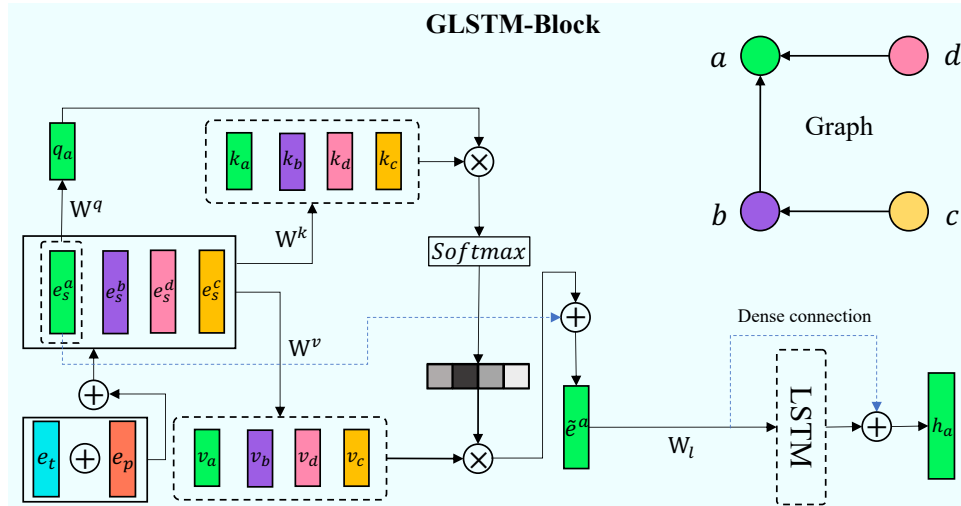


图 4 GLSTM-Block 框架。在本例中，高速公路网络表示为图（Graph）

Fig. 4 The proposed GLSTM-Block framework. In this example, the highway road network is represented as a graph

在高速路网中，每条道路的通行速度受全局道路通行速度的影响，并且不同时期的影响权重  $w$  不同，定义为动态空间相关性。使用多层 GAN 捕捉高速公路网络的动态空间相关性。第二，对于不同的时间步，同一条道路具有时间相关性，表现出动态连续变化的特性，定义为时间相关性。使用多层 LSTM 来捕捉高速公路网络的动态时间相关性。GLSTM-Block 的工作过程如图 4 所示， $\oplus$  和  $\otimes$  分别表示按元素求和和矩阵乘法。由于本研究将会高频使用非线性变换函数，首先将其定义为：

$$f(x) = \text{ReLU}(xW + b) \quad (1)$$

其中  $x$  代表输入变量， $W$  和  $b$  代表可学习参数，ReLU 代表非线性激活函数。

**动态空间特征提取器** 本文使用 GAN 在每个时间步  $t_i$  提取高速公路网络图中任意节点  $v_j$  的动态空间相关性，如图 4 和图 5 所示。

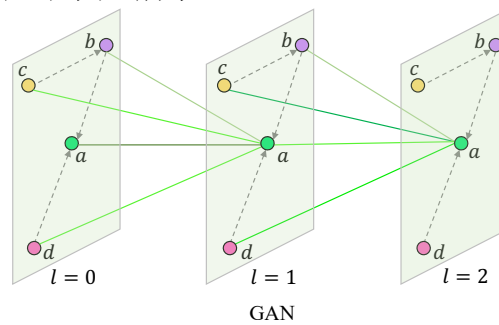


图 5 动态空间关联特征提取过程。图例使用 2 层 GAN 提取高速公路交通数据的动态空间相关性，每层获取全局空间相关性特征

Fig. 5 Dynamic spatial correlation feature extraction process. Example of using a 2-layer GAN to extract the dynamic spatial correlation of the input highway traffic data, where each layer obtains the global correlation features

对于案例图节点 $v_j$ ，输入为 $e_s^{v_j,t_i}$ 和 $e_t^{v_j,t_i}$ ，动态空间相关性可以通过多头自注意力机制计算得到。所有节点 $V$ 对节点 $v_j$ 的影响 $\alpha_{v_j,v}^k$ 可通过第 $k$ 头自注意力计算得到，如下公式所示：

$$\alpha_{v_j,v}^k = \frac{\exp(s_{v_j,v}^k)}{\exp(\sum_{v_r \in V} s_{v_j,v_r}^k)} \quad (2)$$

其中 $s_{v_j,v}^k$ 表示节点 $v_j$ 与节点 $v$ 之间的相关性。

相关性 $s_{v_j,v}^k$ 可以通过节点 $v_j$ 的查询向量和 $v$ 的键向量的内积得到，如下公式所示：

$$s_{v_j,v}^k = \frac{\langle f_k^q((e_s^{v_j,t_i} + e_t^{v_j,t_i})W_k^q), f_k^k((e_s^{v,t_i} + e_t^{v,t_i})W_k^k) \rangle}{\sqrt{d}} \quad (3)$$

其中， $f_k^q$ 和 $f_k^k$ 是查询向量和键向量的非线性转换函数， $\langle, \rangle$ 表示内积操作。

在获得 $\alpha_{v_j,v}^k$ 影响权重后，节点 $v_j$ 的隐藏状态 $\hat{e}_{s,k}^{v_j,t_i}$ 可以通过以下公式计算：

$$\hat{e}_{s,k}^{v_j,t_i} = \sum_{v_r \in V} \alpha_{v_j,v_r}^k V_{v_r}^k + e_s^{v_j} \quad (4)$$

$$V_{v_r}^k = f_k^v((e_s^{v_r,t_i} + e_t^{v_r,t_i})W_k^v) \quad (5)$$

其中 $f_k^v$ 是值向量的非线性转换函数。

节点 $v_j$ 的动态空间相关性通过公式（4）计算后，多头注意力被使用。多头注意力允许模型关注来自不同高速公路不同子空间的信息。尤其是，并行 $M$ 头注意力机制和连接不同的学习表示， $v_j$ 新的隐藏状态可以通过公式（6）更新得到。

$$\hat{e}_s^{v_j,t_i} = [\hat{e}_{s,1}^{v_j,t_i}, \dots, \hat{e}_{s,M}^{v_j,t_i}] W^o \quad (6)$$

其中 $W^o \in R^{N \times d \times d}$ 表示映射层，它可以帮助模型整合高速公路 $v_j$ 所有相关方面的空间相关性。

在 $t_i$ 时间步，每条路 $v_j$ 的空间相关性可以通过公式（1）到（6）计算得到，并且整个路网的输出为 $x_{t_i}^* \in R^{N \times d}$ ，如下公式所示：

$$x_{t_i}^* = \hat{e}_s^{t_i} W_l \quad (7)$$

其中， $\hat{e}_s^{t_i}$ 表示动态空间相关性提取器的输出， $W_l \in R^{d \times d}$ 。

**动态时间特征提取器** 时间相关性是高速公路交通数据的重要特征之一，以往的研究主要从时间维度进行研究<sup>[11, 17]</sup>。LSTM是时间特征提取的主流方法，本文将用作时间特征提取器。多层 LSTM 被用来来捕获公路交通数据的时间相关性。时间特征提取器的工作过程如图 6 所示。

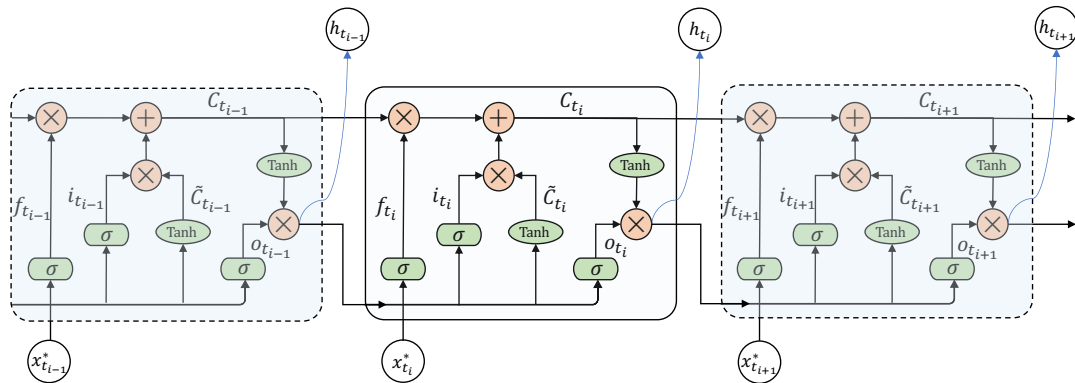


图 6 LSTM 模型结构



Fig. 6 LSTM model structure

对于节点 $v_j$ ，空间特征提取器的输出为 $x_{v_j,t_i}^* \in R^d$ ，使用 LSTM 提取其时间特征。假设 $i$ 、 $f$ 、 $o$ 分别代表输入门、遗忘门和输出门， $x_{v_j,t_i}^*$ 是每个时间步 $t_i$ 的输入， $h_{v_j,t_i}$ 是在每个时间步 $t_i$ 的输出。 $\sigma$ 代表 sigmoid 函数， $\tanh$ 代表 Tanh 函数， $W$ 和 $b$ 分别代表 LSTM 的权重和偏置参数。时间相关性的提取过程如下面的等式所示：

步骤 1：LSTM 选择性地忘记时间步 $t_i$ 时细胞状态 $c_{v_j,t_{i-1}}$ 的特征信息：

$$f_{v_j,t_i} = \sigma(W_f x_{v_j,t_i}^* + W_f h_{v_j,t_{i-1}} + b_f) \quad (8)$$

$$c'_{v_j,t_i} = f_{v_j,t_i} \circ c_{v_j,t_{i-1}} \quad (9)$$

步骤 2：LSTM 从输入特征中选取重要信息，用于更新状态单元 $c'_{v_j,t_i}$ ：

$$\tilde{c}_{v_j,t_i} = \tanh(W_c h_{v_j,t_{i-1}} + W_c x_{v_j,t_i}^* + b_c) \quad (10)$$

$$i_{v_j,t_i} = \sigma(W_i h_{v_j,t_{i-1}} + W_i x_{v_j,t_i}^* + b_i) \quad (11)$$

$$c_{v_j,t_i} = c'_{v_j,t_i} + i_{v_j,t_i} \circ \tilde{c}_{v_j,t_i} \quad (12)$$

步骤 3：最后确定 LSTM 输出：

$$o_{v_j,t_i} = \sigma(W_o h_{v_j,t_{i-1}} + W_o x_{v_j,t_i}^* + b_o) \quad (13)$$

$$h_{v_j,t_i} = o_{v_j,t_i} \circ \tanh(c_{v_j,t_i}) \quad (14)$$

上述整个 GLSTM-Block 工作流程，介绍了在时间步 $t_i$ 高速公路交通数据时空特征 $h_{t_i} \in R^{N \times d}$ 的提取过程，包括动态空间特征和时间特征。

## 2.5 编码器

数据的时空特征提取过程已经在 2.4 部分详细介绍。为了防止每条高速道路过分依赖周边道路信息，并保持自身个性化的时间序列特征。本研究独立于 GLSTM-Block，专门使用时间序列网络 LSTM 来提取原始输入数据的时间特征，如图 3 中编码器部分所示。LSTM 的时间序列特征提取过程如公式（8）到（14）所示，并且输出时间特征为 $h^T = \{h_{t_1}^T, \dots, h_{t_P}^T\}$ ， $h_{t_i}^T \in R^{N \times d}$ 。最后，将 LSTM 的输出和 GLSTM-Block 的输出 $h^{ST} = \{h_{t_1}^{ST}, \dots, h_{t_P}^{ST}\}$ ， $h_{t_i}^{ST} \in R^{N \times d}$ 进行相加，得到最终的时空特征 $\bar{h} = \{\bar{h}_{t_1}, \dots, \bar{h}_{t_P}\}$ ， $\bar{h}_{t_i} = (h_{t_i}^T + h_{t_i}^{ST}) \in R^{N \times d}$ 。

给定输入序列 $x = \{x_{t_1}, \dots, x_{t_P}\}$ ， $x_{t_i} \in R^{N \times d}$ ，由 GLSTM-Block 和 LSTM 构成的编码器用于将输入数据的时空特征转换为隐藏表示向量 $\bar{h} = \{\bar{h}_{t_1}, \dots, \bar{h}_{t_P}\}$ ， $\bar{h}_{t_i} \in R^{N \times d}$ ，稍后用于解码器的时间注意力层。

## 2.6 解码器

解码器内部组成类似于编码器，基于 GLSTM-Block，GLSTM-Block 输出为 $\tilde{h} = \{h_{t_{P+1}}^{ST}, \dots, h_{t_{P+Q}}^{ST}\}$ ， $h_{t_{P+i}}^{ST} \in R^{N \times d}$ 。不同之处在于解码器还包含其他两个层：一个时间注意层和一个子任务层。时间注意力层使用多头自注意力机制来关注历史高速公路交通数据的时空特征对每个时间步预测值的影响<sup>[33]</sup>，子任务层用全连接层实现高速公路速度预测。

### 2.6.1 时间注意力

直到时间步 $t_{P+i}$ ，编码器和解码器输出的时空特征分别为 $\bar{h}$ 和 $\tilde{h}$ ，表示为 $\hat{h} = \bar{h} \cup \tilde{h} = \{\bar{h}_{t_1}, \dots, \bar{h}_{t_P}, \dots, \tilde{h}_{t_{P+i}}\}$ 。对于示例图节点 $v_j$ ，可以通过多头自注意力机制计算时间相关性<sup>[33]</sup>。从 $t_1$ 到 $t_{P+i}$ 的所有时间步的影响权重 $\beta_{t_{P+i}}^k$ 对时间步 $t_{r+i}$ 的影响权重通过多头自注意力机制计算，如下式所示：

$$\beta_{t_{p+i},t}^k = \frac{\exp(u_{t_{p+i},t})}{\exp(\sum_{t_r \in \mathcal{N}_{t_{p+i}}} u_{t_{p+i},t_r})} \quad (15)$$

其中 $\mathcal{N}_{t_{p+i}}$ 表示 $t_{p+i}$ 之前的一组时间步长， $u_{t_{p+i},t}$ 表示时间步 $t_{p+i}$ 和 $t$ 的时空特征之间的相关性，可以通过节点 $v_j$ 在时间步 $t_{p+i}$ 的查询向量与节点 $v_j$ 在时间步 $t$ 的键向量的内积得到，如以下等式所示：

$$u_{t_{p+i},t} = \frac{\langle f_k^q(\hat{h}_{v_j,t_{p+i}} + e_t^{v_j,t_{p+i}}) W_k^q, f_k^k(\hat{h}_{v_j,t} + e_t^{v_j,t}) W_k^k \rangle}{\sqrt{d}} \quad (16)$$

获得影响权重 $\beta_{t_{p+i},t}^k$ 后，隐藏状态 $\bar{h}_{v_j,t_{p+i}}^k$ 通过等式（16）更新。

$$\bar{h}_{v_j,t_{p+i}}^k = \sum_{t_r \in \mathcal{N}_{t_{p+i}}} \beta_{t_{p+i},t_r}^k V_{t_r}^k + \hat{h}_{v_j,t_{p+i}} \quad (17)$$

$$V_{t_r}^k = f_k^v(\hat{h}_{v_j,t_r} + e_t^{v_j,t_r}) W_k^v \quad (18)$$

$v_j$ 的时间相关性由影响权重 $\beta_{t_{p+i},t}^k$ 和 $V^k = \{V_t^k | t \in \mathcal{N}_{t_{p+i}}\}$ 相乘得到，多头自注意力机制被采用。

$$h_{t_{p+i}}'' = [\bar{h}_{v_j,t_{p+i}}^1, \dots, \bar{h}_{v_j,t_{p+i}}^M] W^o \quad (19)$$

在时间步 $t_{p+i}$ ，时间相关性可以通过上述（15）到（19）式计算得到，整个高速公路网络的输出为 $h_{t_{p+i}}'' \in R^{N \times d}$ 。

### 2.6.2 高速公路交通速度预测

对于高速公路交通速度预测任务，将解码器的输出特征直接馈送到全连接层以生成预测值。

$$y = h'' W_s \quad (20)$$

其中 $W_s \in R^{d \times 1}$ 表示全连接层的权重参数。

### 2.6.3 损失函数

目标损失函数，选择预测值 $\hat{y}$ 和观测值 $y$ 之间的平均绝对误差（MAE）函数作为交通速度预测任务的损失函数。

$$loss = \frac{1}{Q \times N} \sum_{v_r}^V \sum_{t=t_{p+1}}^{t_{p+Q}} |y_{v_r,t} - \hat{y}_{v_r,t}| \quad (21)$$

## 3. 实验

### 3.1 基线模型和评价方法

对于高速公路交通速度预测，将本文提出的 ST-ANet 模型与以下基线方法进行对比：**HA** 模型在未来预测任务中使用每天同一时间的历史数据的平均值作为预测值<sup>[3]</sup>。**ARIMA** 是一种传统的时间序列预测方法，它结合了移动平均和自回归分量对历史时间序列数据建模<sup>[5]</sup>。**SVM** 是一种称为支持向量回归技术，用于交通速度的短期预测<sup>[6]</sup>。**Bi-LSTM NN** 由双向长短期记忆神经网络构成，用于对每个关键路径进行建模，然后使用堆叠在一起的多个 Bi-LSTM 层来合并时间序列信息<sup>[18]</sup>。**FI-RNNs** 将特征注入循环神经网络，结合时间序列数据并使用堆叠的 RNN 和编码器来学习交通数据的序列特征<sup>[19]</sup>。**HyperNet** 使用 LSTM 和 RNN 网络来预测交通速度，使用自动机器学习方法搜索模型超参数以执行网络训练<sup>[20]</sup>。**Multi-view CNN** 基于传统 CNN 网络，将历史和实时交通数据作为模型输入，以预测交通速度<sup>[21]</sup>。**PSPNN** 基于路径的速度预测神经网络，由 CNN 和双向 LSTM (Bi-LSTM) 网络组成，用于提取历史数据的时空特征，实现基于路径的速度预测<sup>[24]</sup>。**MDL** 是一种新颖的混合深度学习模型，用于预测车道级短期交



通速度，它由卷积长短期记忆 (Conv-LSTM) 层、卷积层和全连接层组成<sup>[26]</sup>。**OGCRNN** 基于 GCN 和 GRU 网络的时空神经网络模型，用于提取路段之间交通演变的复杂时空特征以预测交通速度<sup>[30]</sup>。**GCN-LSTM** 基于 GCN 和 LSTM 网络的融合深度模型，用于提取路段之间交通数据复杂的时空特征以预测交通速度<sup>[31]</sup>。

为了评估 ST-ANet 模型的预测性能，将使用三个评价指标来评估观测值  $y_i$  和预测值  $\hat{y}_i$  之间的差异：均方根误差 (RMSE)、平均绝对误差 (MAE) 和决定系数 ( $R^2$ )。评价指标公式如下：

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{1}{T} \sum_{i=1}^T (y_i - \hat{y}_i)^2} \quad (22)$$

$$\text{MAE}(y, \hat{y}) = \frac{1}{T} \sum_{i=1}^T |y_i - \hat{y}_i| \quad (23)$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (24)$$

其中  $T$  是测试集大小， $\bar{y}$  表示的是观测值的平均值。高  $R^2$  和低 RMSE 和 MAE 值表明预测性能更精准。

### 3.2 模型参数

ST-ANet 模型中的超参数在训练过程中确定，即通过 MAE 在测试集上选择性能最好的模型。首先，根据经验手动设定超参数范围：学习率 {0.01, 0.005, 0.001, 0.0005}，dropout {0.0, 0.1, 0.2, 0.3, 0.4, 0.5}，正则化参数 {0.1, 0.01, 0.001, 0.0001}，衰减率 {0.99, 0.95, 0.90, 0.85}。对于 ST-ANet 模型，发现以下设置效果最好：将 dropout 设置为 0.5，衰减率设置为 0.99，正则化参数设置为 0.0001，学习率设置为 0.0005。使用基线模型时，这些设置仍然有效。所有的实验模型通过 Tensorflow 框架实现，并使用 Batch 大小为 32 的随机梯度下降算法 (SGD) 优化器训练所有模型。在所有实验中，模型使用了提前停止机制，即提前停止轮次和最大 epoch 分别设置为 20 和 50。

经过多次训练，最终确定了模型框架参数。如表 1 所示，已经详细列出了 ST-ANet 模型的层数、节点数、输出大小和相关超参数。本文提出的 ST-ANet 模型和对比的基线模型具体实现代码已经提交到代码开源社区平台 GitHub，请参考个人 GitHub 主页 (<https://github.com/zouguojian/Traffic-speed-prediction/tree/main/ST-ANet>)。

表 1 模型参数

层名	输出形状	参数	值
GAN	64	$q$ : 节点×层数 $k$ : 节点×层数 $v$ : 节点×层数 相同层数 M	64×1 64×1 64×1 4 1
LSTM	64	节点×层数	64×1
Temporal attention	64	$q$ : 节点×层数 $k$ : 节点×层数 $v$ : 节点×层数 相同层数 M	64×1 64×1 64×1 4 1
Dense connection	64	节点×层数	64×1
Feed layer	64	节点×层数	64×1

Full connected layer	128 64 1	节点×层数	128×1 64×1 1×1
-	-	Batch	32
-	-	Dropout	0.5
-	-	衰减率	0.99
-	-	学习率	0.0005
-	-	Epoch	50
-	-	$\lambda$	0.0001
-	-	训练方法	SGD
-	-	$\gamma$	0.87

### 3.2 模型表现对比

#### 3.2.1 单步预测表现对比

交通速度单步预测任务意义显著，也是目前研究的热点话题。大多数交通速度预测研究集中在下一时刻的单步预测，这与区域高速路网的交通运行状况紧密相关。因此，本文继续将下一时刻的交通速度预测作为研究目标任务，预测结果如表 2 所示。实验中，使用 6 小时的历史数据来预测下一小时的交通速度。例如，上午 5:00-11:00 是输入时段，上午 11:00-12:00 为预测时段。

表 2 [6-1 h] 任务，通过不同方法获得的交通速度预测结果

模型名称	MAE	RMSE	R <sup>2</sup>
HA <sup>[3]</sup>	14.360	22.867	0.427
ARIMA <sup>[5]</sup>	15.896	23.137	0.269
SVM <sup>[6]</sup>	14.342	23.095	0.603
Bi-LSTM NN <sup>[18]</sup>	13.912	23.714	0.583
FI-RNNs <sup>[19]</sup>	13.760	23.721	0.583
HyperNet <sup>[20]</sup>	13.875	23.714	0.583
Multi-view CNN <sup>[21]</sup>	14.205	22.274	0.632
PSPNN <sup>[24]</sup>	14.293	21.927	0.643
MDL <sup>[26]</sup>	13.630	21.626	0.653
OGCRNN <sup>[30]</sup>	13.360	21.547	0.656
GCN-LSTM <sup>[31]</sup>	12.985	<b>21.383</b>	<b>0.661</b>
GLSTM-Block	12.735	21.438	0.657
GLSTM-Block-LSTM	12.614	21.418	0.658
GLSTM-Block-Temporal-Attention	12.696	21.702	0.649
ST-ANet	<b>12.470</b>	21.460	0.657

表 2 显示本文提出的 ST-ANet 和基线模型在下一小时的高速公路交通速度预测任务中的结果。通过对比，HA 和 ARIMA 的性能远不如所有其他基线模型，这说明了高速公路交通速度预测的难度。统计模型最佳 MAE、RMSE 和 R<sup>2</sup> 值分别为 14.360、22.867 和 0.427。与统计方法相比，传统机器学习方法和深度学习方法表现更好，因为它们更适合提取时空数据的非线性特征。因此，现有的研究正在逐渐从统计方法过渡到机器学习方法。

将基于 RNN 和 CNN 的基线模型分为时间依赖、空间依赖和时空依赖三大类。时间依赖模型包括 Bi-LSTM NN、FI-RNNs 和 HyperNet，空间依赖模型包括 Multi-view CNN，时空依赖模型包括 PSPNN 和 MDL。比较这些基线模型，可以发现一些有价值的结果：首先，时间依赖模型的整体预测误差低于空间依赖模型，最佳 MAE 值为 13.760。结果表明，时间特征显著影响交通速度预测的精准性，体现使用基于 RNN 网络提取输入数据时间相关性的优势。其次，空间依赖模型的交通速度预测偏差和拟合

优度优于时间依赖模型，最佳 RMSE 和  $R^2$  值分别为 22.274 和 0.632。基于 CNN 的空间依赖模型的实验结果表明，空间相关性在交通速度预测中起到了积极的作用，有效缓解了预测偏差过大和拟合优度低的问题。第三，与时间依赖模型和空间依赖模型相比，时空依赖模型提高了交通速度预测的性能显著。最佳 MAE、RMSE 和  $R^2$  值分别为 13.630、21.626 和 0.653。实验结果表明，结合输入数据的时间相关性和空间相关性两个方面的特征，可以有效提高公路交通速度预测的准确性。

在基线模型中，基于 GCN 和 RNN 的时空依赖模型，包括 OGCRNN 和 GCN-LSTM。对比 OGCRNN，GCN-LSTM 和 PSPNN 模型，OGCRNN 和 GCN-LSTM 的预测表现更好，并且对应的最优 MAE，RMSE 和  $R^2$  的值分别为 12.985, 21.383 和 0.661。实验证明 GCN 解决了 CNN 网络面临的问题，可以有效地提取非欧式空间中离散道路之间交通数据的空间相关性。对比 OGCRNN 和 GCN-LSTM 的预测结果，GCN-LSTM 的预测结果明显好于 OGCRNN。GCN-LSTM 的预测结果证明，LSTM 相对于 GRU 更适用于时间序列特征提取。本研究继续考虑非欧式空间中交通数据的空间离散分布问题，将 GCN 延伸到 GAN 模型：第一，使用基于空间注意力的 GAN 提取高速公路网络动态空间相关性。第二，与 GCN-LSTM 模型类似，使用 LSTM 作为时间特征提取器获取输入数据的时间相关性。第三，使用时间注意力机制计算历史输入数据与当前预测值之间的相关性，以完成交通速度预测。通过将本文提出的 ST-ANet 模型与基线模型对比，ST-ANet 模型的预测性能得到较大的提升。与最优基线模型 GCN-LSTM 相比，模型的预测误差改善了 4.0%。

### 3.2.2 模型每个部分贡献

**GLSTM-Block:** 与 GCN-LSTM 模型相比，相同之处在于，他们都是使用 LSTM 作为时间特征提取器；不同之处是，GLSTM-Block 使用基于空间注意力机制的 GAN 提取输入数据的动态空间相关性，而 GCN-LSTM 是依托于 GCN 提取静态空间相关性。GLSTM-Block 与 GCN-LSTM 相比，GLSTM-Block 取得了更好的表现，误差 MAE 降低了 1.9%。实验结果证明：一方面，相比于 GCN，GAN 考虑到了空间特征的动态变化，这对高速公路网的空间相关性特征提取效果要更加显著；另一方面，考虑路段之间的动态影响可以有效缓解预测能力不足问题。

**GLSTM-Block-LSTM:** 高速路网中的每条公路不仅受周边公路影响，更受到自身速度的影响。相比于 GLSTM-Block，模型在 Encoder 部分添加独立于 GLSTM-Block 模块用于提取每条公路速度时间序列特征的 LSTM 网络，目的是为了保留每个路段自身的时序特征而不受周边路段的影响。与 GLSTM-Block 预测结果相比，GLSTM-Block-LSTM 模型的预测表现得到提升，并且 MAE、RMSE 和  $R^2$  三个评判指标测量值都得到更新，分别改善了 1.0%、0.1% 和 0.2%。通过实验，可以得到道路的自身固有时序特征对道路交通速度预测任务影响较大。因此，在高速公路速度预测任务中，不仅需要考虑到高速路网时空特征问题，还需要保留每一条道路固有的时间序列特征。

**GLSTM-Block-Temporal-Attention:** 高速公路未来时间段的交通速度预测值会受到过去一段时间的交通速度影响。为此，本文提出了基于时间注意力的 GLSTM-Block-Temporal-Attention 模型。与 GLSTM-Block 相比，GLSTM-Block-Temporal-Attention 在预测阶段会考虑当前预测与历史输入数据之间的相关性。通过与 GLSTM-Block 实验对比，可以得到添加时间注意力后的模型预测表现整体误差减小，误差 MAE 值降低了 0.3%。模型的误差降低的可能原因是，更加关注时间维度历史输入数据与预测值之间的相关性。这也反映了时间注意力机制在交通预测任务中的积极作用，可作为交通速度预测模型的重要组件之一。

ST-ANet：通过对 GLSTM-Block，GLSTM-Block-LSTM 和 GLSTM-Block-Temporal-Attention 模型预测结果的分析，可总结得到基于空间注意力的 GAN、LSTM 和时间注意力机制对交通速度的预测具有不同程度的影响。为了综合三者的优势，本研究向 GLSTM-Block 模型当中添加用于独立提取高速公路每条道路交通速度时间序列特征的 LSTM，以及用于计算历史输入数据与当前预测值相关性的时间注意力机制，最终得到 ST-ANet 模型。与最优基线模型 GCN-LSTM 相比，ST-ANet 的预测误差得到了大幅度改善，误差 MAE 降低了 4.0%。

### 3.2.3 长期预测表现对比

目前的交通速度预测研究主要集中在短期预测，不足以满足实际应用场景的需要。高速公路交通速度的长期预测对于预测模型来说是一项具有挑战性的任务，它关系到未来一段时间公路交通的精确管控。对于表 2 中的结果，已经证明了 ST-ANet 在短期预测任务中的有效性，并分析了模型各部分的贡献。为了验证 ST-ANet 在长期预测任务中的优势，表 3 分别显示了基线模型和提出的 ST-ANet 模型在未来 2 小时和 3 小时的交通速度预测任务中的表现。由于基线模型的局限性，本阶段从表 2 中选择性能好，并适用于长期预测的最先进基线模型进行实验对比。对于[6-2 h]交通速度预测任务，使用 6 个小时的历史交通数据作为输入，预测未来 2 小时的交通速度。例如 5:00-11:00 为输入时段，11:00-13:00 为预测时段。对于[6-3 h]交通速度预测任务，使用 6 个小时的历史交通数据作为输入，预测未来 3 小时的交通速度。例如 5:00-11:00 为输入时段，11:00-14:00 为预测时段。

表 3 任务[6-2 h]和[6-3 h]，通过不同方法获得的交通速度预测结果

模型名称	[6-2 h] 任务			[6-3 h] 任务		
	MAE	RMSE	R <sup>2</sup>	MAE	RMSE	R <sup>2</sup>
Bi-LSTM NN <sup>[18]</sup>	14.685	24.421	0.558	15.562	25.499	0.518
FI-RNNs <sup>[19]</sup>	14.573	24.296	0.562	15.355	25.404	0.521
HyperNet <sup>[20]</sup>	14.738	24.478	0.556	15.409	25.422	0.521
Multi-view CNN <sup>[21]</sup>	14.773	22.783	0.615	15.217	23.348	0.596
PSPNN <sup>[24]</sup>	15.186	22.630	0.620	15.595	23.139	0.603
MDL <sup>[26]</sup>	14.157	22.320	0.631	14.403	22.440	0.627
OGCRNN <sup>[30]</sup>	13.657	21.931	0.643	13.999	22.306	0.631
GCN-LSTM <sup>[31]</sup>	13.240	21.863	<b>0.646</b>	13.542	<b>21.901</b>	<b>0.644</b>
ST-ANet	<b>12.764</b>	<b>21.787</b>	<b>0.646</b>	<b>13.009</b>	22.187	0.633

为了更清楚地对不同模型的预测结果进行对比，图 7 (a)、图 7 (b)和图 7(c)对每个模型的预测误差、偏差和拟合优度进行了可视化。结合图 7 和表 3，将 ST-ANet 模型与用于长期预测任务的基线模型相比，ST-ANet 表现最佳。对于交通预测任务[6-2 h]，与最优基线模型 GCN-LSTM 相比，MAE 改进为 3.6%，RMSE 改进为 0.3%。对于交通预测任务[6-3h]，与基线模型相比，MAE 改进为 3.9%。此外，如图 7 所示，与基线模型相比，ST-ANet 的误差随着时间步数的增加变动幅度较小，且误差小于所有基线模型。ST-ANet 的预测结果优势显著，因此具有巨大的应用前景。值得注意的是，ST-ANet 的预测偏差和拟合优度要略低于 GCN-LSTM，原因可能是 GCN 对于降低模型偏差和提高拟合优度方面的作用要好于 GAN，而 GAN 对于降低模型的误差方面起着不可或缺的作用。ST-ANet 和 GCN-LSTM 模型的特性给未来的交通路网交通速度预测带来了新的灵感，深度结合 GCN 和 GAN 的优势是本研究未来研究的一个重要方向。

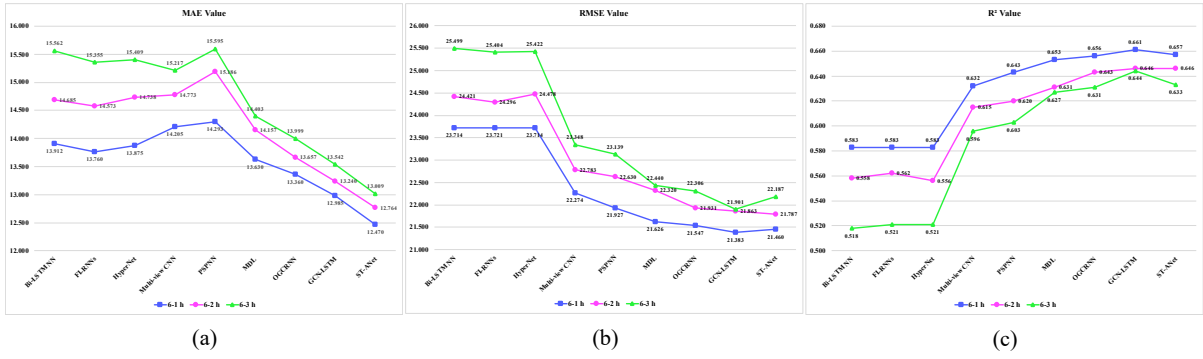


图 7 长期交通速度预测能力。(a) 在交通速度预测任务中, 不同模型 MAE 的变化; (b) RMSE 的变化; (c)  $R^2$  的变化

Fig. 7 Long-term traffic speed prediction ability. (a) MEA changes of different models in traffic speed prediction tasks; (b) RMSE changes; (c)  $R^2$  changes

#### 4. 结论

本文提出了一种面向动态高速路网的时空注意力网络速度预测模型 ST-ANet, 可以精准预测高速公路交通速度。首先, 使用基于多头自注意力的 GAN 和 LSTM 组成的 GLSTM-Block 提取高速公路交通数据的动态空间相关性和时间相关性。其次, 考虑到每条道路的独特性, 使用基于 LSTM 的时间序列特征提取器提取高速公路每条道路自身的时间相关性, 保留每条道路自身的上下文语义信息。最后, 还使用基于多头自注意力的时间注意力机制计算每个时间步的历史输入数据和预测值之间的相关性。此外, 为了防止网络特征丢失和内部协变量偏移, 还为每个网络层添加了密集连接和层归一化技巧。

为了评估所提出模型 ST-ANet 的性能, 对真实世界的高速公路交通数据进行了多次实验。ST-ANet 使用 GLSTM-Block、LSTM 和时间注意力机制有效地提取公路交通数据的时空特征。对于高速公路交通速度短时预测, 在相同的测试集上, 所提出的 ST-ANet 比基线方法具有更准确的交通速度预测性能。对比最优基线模型 GCN-LSTM, ST-ANet 模型对于未来 1 个小时预测误差 MAE 降低了 4.0%。此外, 对于长期预测任务, ST-ANet 的预测误差、偏差和拟合优度的变化范围很小。对比最优基线模型 GCN-LSTM, ST-ANet 模型对于未来 2 个小时和 3 个小时内高速路网交通速度的预测误差 MAE 分别降低了 3.6%和 3.9%, 这也证明了其巨大的应用价值。在未来的工作中, 本文希望在提出的 ST-ANet 模型中添加 GCN 网络, 用于减小模型在交通速度预测任务中的偏差和提高模型的拟合优度, 并探索用于未来预测的高级模型。

#### 参考文献:

#### References:

- [1] GUO F, POLAK J W, KRISHNAN R. Predictor fusion for short-term traffic forecasting[J]. Transportation Research Part C: Emerging Technologies, 2018, 92(7): 90-100.
- [2] 林培群, 夏雨, 周楚昊. 引入时空特征的高速公路行程时间预测方法[J]. 华南理工大学学报(自然科学版), 2021, 49(8): 1-11.  
LIN P Q, XIA Y, ZHOU C H. Freeway travel time prediction method with spatio-temporal characteristics[J]. Journal of South China University of Technology (Natural Science Edition), 2021, 49(8): 1-11. (in Chinese)
- [3] LIU L B, QIU Z, LI G B, et al. Contextualized spatial-temporal network for taxi origin-destination demand prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(10): 3875-3887.

- [4] 谢海红, 戴许昊, 齐远. 短时交通流预测的改进 K 近邻算法[J]. 交通运输工程学报, 2014, 14(3): 87-94.  
XIE H H, DAI X H, QI Y. Improved K-nearest neighbor algorithm for short-term traffic flow forecasting[J]. Journal of Traffic and Transportation Engineering, 2014, 14(3): 87-94. (in Chinese)
- [5] WU W, XIA Y, JIN W. Predicting bus passenger flow and prioritizing influential factors using multi-source data: scaled stacking gradient boosting decision trees[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(4): 2510-2523.
- [6] VANAJAKSHI L, RILETT L R. A comparison of the performance of artificial neural networks and support vector machines for the prediction of traffic speed[C]// Proceeding of the IEEE Intelligent Vehicles Symposium. Parma, ITALY: IEEE, 2004: 194-199.
- [7] JIANG B N, FEI Y S. Vehicle speed prediction by two-level data driven models in vehicular networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(7): 1793-1801.
- [8] SHIN J, SUNWOO M. Vehicle speed prediction using a Markov chain with speed constraints[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 20(9): 3201-3211.
- [9] ZHANG W B, FENG Y Y, LU K, et al. Speed prediction based on a traffic factor state network model[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(5): 3112-3122.
- [10] 张文胜, 郝孜奇, 朱冀军, 等. 基于改进灰狼算法优化 BP 神经网络的短时交通流预测模型[J]. 交通运输系统工程与信息, 2020, 20(2): 196-203.  
ZHANG W S, HAO Z Q, ZHU J J, et al. BP neural network model for short-time traffic flow forecasting based on transformed Grey wolf optimizer algorithm[J]. Journal of Transportation Systems Engineering and Information Technology, 2020, 20(2): 196-203. (in Chinese)
- [11] CSIKOS A, VIHAROS Z J, KIS K B, et al. Traffic speed prediction method for urban networks - an ANN approach.[C]// Proceedings of the International Conference on Models & Technologies for Intelligent Transportation Systems. Budapest, HUNGARY: IEEE, 2015: 102-108.
- [12] JIA Y H, WU J P, DU Y M. Traffic speed prediction using deep learning method[C]// Proceedings of the 19<sup>th</sup> IEEE International Conference on Intelligent Transportation Systems (ITSC). Rio de Janeiro, BRAZIL: IEEE, 2016: 1217-1222.
- [13] TANG J J, LIU F, ZOU Y J, et al. An improved fuzzy neural network for traffic speed prediction considering periodic characteristic[J]. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(9): 2340-2350.
- [14] 李桃迎, 王婷, 张羽琪. 考虑多特征的高速公路交通流预测模型[J]. 交通运输系统工程与信息, 2021, 21(3): 101-111.  
LI T Y, WANG T, ZHANG Y Q. Highway traffic flow prediction model with multi-features[J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21(3): 101-111. (in Chinese)
- [15] ZHANG B, ZOU G J, QIN D M, et al. A novel Encoder-Decoder model based on read-first LSTM for air pollutant prediction[J]. Science of The Total Environment, 2021, 765(3): 144507.
- [16] MA X L, TAO Z M, WANG Y H, 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(5): 187-197.
- [17] GU Y L, LU W Q, QIN L Q, et al. Short-term prediction of lane-level traffic speeds: A fusion deep learning model[J]. Transportation Research Part C: Emerging Technologies, 2019, 106(10): 1-16.
- [18] WANG J W, CHEN R X, HE Z C. Traffic speed prediction for urban transportation network: A path based deep learning approach[J]. Transportation Research Part C: Emerging Technologies, 2019, 100(3): 372-385.
- [19] QU L C, LV J, Li W, et al. Features injected recurrent neural networks for short-term traffic speed prediction[J]. Neurocomputing, 2021, 451(9): 290-304.
- [20] YI H, BUI K H N. An automated hyperparameter search-based deep learning model for highway traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(9): 5486-5495.
- [21] JIA D Y, CHEN H B, ZHENG Z D, et al. An enhanced predictive cruise control system design with data-driven traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2021: DOI: 10.1109/TITS.2021.3076494.
- [22] LV Z J, XU, J J, ZHENG K, et al. LC-RNN: A deep learning model for traffic speed prediction[C]// Proceedings of the 27<sup>th</sup> International Joint Conference on Artificial Intelligence (IJCAI-18). Stockholm, Sweden: International Joint Conferences on Artificial Intelligence, 2018: 3470-3476.
- [23] ZHOU L X, ZHANG S C, YU J R, et al. Spatial-temporal deep tensor neural networks for large-scale urban network speed prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(9): 3718-3729.

- [24] YANG H, LIU C X, ZHU M X, et al. How fast you will drive? predicting speed of customized paths by deep neural network[J]. IEEE Transactions on Intelligent Transportation Systems, 2021: DOI: 10.1109/TITS.2020.3031026.
- [25] ZANG D, LING J W, WEI Z H, et al. Long-term traffic speed prediction based on multiscale spatio-temporal feature learning network[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 20(10): 3700-3709.
- [26] LU W Q, RUI Y K, RAN B. Lane-Level Traffic Speed Forecasting: A Novel Mixed Deep Learning Model[J]. IEEE Transactions on Intelligent Transportation Systems, 2020: DOI: 10.1109/TITS.2020.3038457.
- [27] 曾筠程, 邵敏华, 孙立军, 等. 基于有向图卷积神经网络的交通预测与拥堵管控[J]. 中国公路学报, 2021, 34(12): 239-248.  
ZENG Y C, SHAO M H, SUN L J, et al. Traffic prediction and congestion control based on directed graph convolution neural network[J]. China Journal of Highway and Transport, 2021, 34(12): 239-248. (in Chinese)
- [28] ZHANG Z K, LI Y D, SONG H F, et al. Multiple dynamic graph based traffic speed prediction method[J]. Neurocomputing, 2021, 461(10): 109-117.
- [29] LI D, LASENBY J. Spatiotemporal attention-based graph convolution network for segment-level traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2021: DOI: 10.1109/TITS.2021.3078187.
- [30] GUO K, HU Y L, QIAN Z, et al. Optimized graph convolution recurrent neural network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(2): 1138-1149.
- [31] 陈华伟, 邵毅明, 敖谷昌, 等. 面向在线地图的 GCN-LSTM 神经网络速度预测[J]. 交通运输工程学报, 2021, 21(4): 183-196.  
CHEN H W, SHAO Y M, AO G C, et al. Speed prediction by online map-based GCN-LSTM neural network[J]. Journal of Traffic and Transportation Engineering, 2021, 21(4): 183-196. (in Chinese)
- [32] FANG M Y, TANG L L, YANG X, et al. FTPG: A fine-grained traffic prediction method with graph attention network using big trace data[J]. IEEE Transactions on Intelligent Transportation Systems, 2021: DOI: 10.1109/TITS.2021.3049264.
- [33] VASWANI A, SHAZEER N, PARMAR N, et al. Attention Is All You Need[C]// Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS). Long Beach, CA, USA: Advances in Neural Information Processing Systems, 2017: 6000-6010.

## 备注信息

手机号: 18321682236、邮箱: 2010768@tongji.edu.cn、刊物邮寄地址: 上海市同济大学(嘉定校区)、收件人: 邹国建。