文章编号：0451－0712（20xx）-xxxx-xx 中图分类号：U292.4 文献标识码：A

基于多任务学习的节假日高速公路交通流预测

柴庆刚[[1]](#footnote-1)，章树鑫2\*，邱豪2，吕梦琪3，张云帆4，王骋程3，张金雷2

(1.山东高速基础设施建设有限公司 济南市 250098；2.北京交通大学系统科学学院 北京市 100044；3.山东高速集团有限公司创新研究院 济南市 250098；4. 山东省交通规划设计院集团有限公司 济南市 250098)

**摘 要：**准确的交通流预测可以为高速公路的运营决策提供数据支持，尤其是节假日期间，交通流骤增给高速公路运营管理带来较大压力，运营管理部门需要准确预测未来交通流数据。为此，此研究以假期期间高速公路交通流为研究对象，构建基于双向卷积长短时记忆网络和注意力交互模块的时空注意力交互网络，用以协同预测假期期间高速公路收费站进出站交通流。在山东省五一假期的高速公路收费站进出站交通流数据集上比较模型与其他八个先进基准模型在不同预测步长下的预测性能，其中进站流预测方面相比第二好的模型预测效果平均提升4.34%，出站流预测方面相比第二好的模型预测效果平均提升2.3%，表明模型具有良好的预测性能，可以准确预测节假日期间未来多步的交通流。

**关键词：**公路运输；交通流预测；深度学习；高速公路；交通大数据；时空特征建模

**1 引言**

高速公路作为我国综合交通网络的骨干，在综合运输体系中承担着重要的任务[1]。随着经济社会发展，人们的出行需求逐渐增多，高速公路的车流量随之增大，尤其在特定节假日期间（例如五一、国庆等），国家实施高速公路免费通行政策，使高速公路车流量激增，极易引发道路长时间、长距离拥堵，危害出行者的出行安全[2]。如何充分挖掘节假日期间高速公路交通流特性，实现准确、可靠的交通流预测，对于高速公路的车流管控具有重要意义。

近年来，随着人工智能的飞速发展，国内外越来越多学者利用深度学习网络捕捉交通流的时空特征，从而实现交通流预测。2014年，Ma等[3]第一次将深度学习长短时记忆网络（LSTM）应用至交通流预测领域，取得良好预测效果。随后，学者们提出了大量基于深度学习的交通流预测模型，弥补了深度学习在交通流预测的应用空缺。例如，Zhang[4]等充分考虑了道路交通流的时间邻阶特性、周期特性以及趋势特性，提出了一个全新的时空残差网络（ST-ResNet），准确预测城市道路的交通流。Li[5]等将交通流视为有向图结构上的扩散过程，并提出扩散卷积循环神经网络（DCRNN）实现交通流预测。考虑到交通网络本身是一个具有拓扑信息的图结构，许多学者将图卷积神经网络（GCN）聚合到预测模型中，有效捕捉客流的空间依赖性。例如，Zhao等人[6]利用图卷积网络(GCN)建模交通流网络中复杂的空间相关性，利用门控递归单元(GRU)捕捉动态时间依赖性，从而实现精确的客流预测。Zhang[7]等提出一个图神经网络模型Conv-GCN，利用多图卷积捕捉客流的空间特性，并从实时、日时和周时等三个角度学习客流的时间特性。然而，传统的GCN通过预定义的图建模空间特性，忽略了空间特征的复杂动态性，为此部分学者对GCN网络进行改进，旨在充分捕捉复杂时变的空间特性。例如，Wu[8]等提出一个自适应的依赖性矩阵用以自动挖掘交通流的隐藏的空间依赖性。Zhang[9]等考虑了三种不同类型的图用以建模复杂的客流空间依赖性，分别是邻阶矩阵图、功能相似性图以及OD关联性图；同时引入了一个自适应邻阶矩阵用以刻画动态的空间相关性。注意力机制[10]可以有效捕捉两个节点之间的相关性而不用考虑节点的距离，可以有效解决时间序列数据的长程依赖性问题，因此近年来也被广泛应用于时间序列数据处理中[11-13]。例如，Yan[14]等提出一个交通Transformer，模型包含了一个全局编码器和一个全局-局部编码器，充分利用了注意力机制学习交通流数据中动态多层的时空特性。Xu[15]等提出一个全新的时空Transformer网络（STTNs），模型通过多头注意力机制捕捉动态有向的空间依赖性和长时时间依赖性，实现长时交通流预测。Zhang[16]等利用注意力机制挖掘实时、日时以及周时需求间的时间交互信息，有效建模了交通需求复杂的时间特性。

虽然上述研究有效提升了交通流的预测精度，但是大部分研究均是面向城市道路交通流或城市轨道交通客流，较少研究是面向高速高路的交通流；同时，大部分研究针对常态场景下交通流展开预测，对于节假日等非常态场景下的交通流预测研究较少。事实上，节假日期间的交通流具有极大不确定性和突发性，容易造成交通流拥堵，甚至危险生命财产安全[17]。另外，大部分模型仅考虑短时预测，预测总时间跨度较小。在节假日等持续时间较长的非常规场景下，交通管理者更希望掌握未来较长一段时间（一天以上）的准确的交通流信息，从而制定更鲁棒的管控策略。因此，本研究以节假日期间的高速公路交通流为例，研究非常态场景下的高速公路长时交通流预测问题。

为了解决上述存在的问题，本文提出了一个基于多任务学习的进出站交通流协同预测模型，旨在实现假期期间的高速公路收费站进出站交通流的协同多步预测。由于高速公路收费站的进出站交通流间存在相互影响，将进站交通流和出站交通流单独建模无法充分学习和利用二者间的关联性。因此本研究采用多任务学习训练方式，通过共享二者特征信息的方式处理进站交通流和出站交通流时间序列，利用进站流特征辅助出站流预测，利用出站流特征辅助进站流预测，最终提升模型在进出站流协同预测上的精度。为了充分挖掘进站和出站交通流的时空依赖性，本文提出一个时空注意力交互网络（ST-Cross-Attn），该网络由双向卷积长短时记忆网络（Bi-ConvLSTM）和交互注意力（Cross-Attention）构成。前者用于分别挖掘进站交通流和出站交通流的时空依赖性，后者则将进站和出站交通流的时空依赖性进行交互融合，从而获得实现进站交通流和出站交通流的时空信息共享，获得增强的时空特征。最后，基于构建的时空注意力交互网络，本文搭建了序列到序列（Seq2Seq）的多步模型预测框架，利用编码器获得当前时刻进出站交通流的增强特征编码，利用解码器依次迭代预测得到未来特定个时间步的进出站交通流。文章将预测模型应用于山东省高速公路五一假期期间收费站交通流的真实数据集中，预测实验结果表明本文提出的模型具有较高的预测精度，可以准确预测节假日期间高速公路收费站的进出站交通流。本文的创新点总结如下：

(1) 本文使用多任务学习的训练方式实现高速公路进站交通流和出站交通流的协同预测，相较于独立预测收费站的进站交通流和出站交通流，本文采用的多任务学习方式可以实现进站交通流和出站交通流间时空分布信息的共享，使二者可以利用对方的特征信息辅助预测未来交通流，更利于交通流数据的时空特征挖掘。

(2) 本文构建基于双向卷积长短时记忆网络以及交互注意力的时空注意力交互网络，对进站和出站交通流数据的时空依赖性进行挖掘，并实现二者间时空特征信息的交互，有利于增强数据中的时空特征并提高预测精度。

(3) 本文采用“序列到序列”的多步预测框架，通过编码器逐步编码获得当前时刻的隐藏特征，并通过解码器逐步迭代实现多步预测。实验结果表明，在使用该框架进行多步预测时，模型的累积误差较小，依然能够保持较高的预测精度。

**2 问题描述与建模**

2.1 问题描述

本研究的目标是基于历史假期前后的高速公路收费站交通流数据，协同预测未来假期若干个时间步内的高速公路收费站进出站交通流流量。给定历史假期的收费站进出站交通流数据，本研究选取60分钟作为时间粒度，从中提取收费站进站交通流时间序列数据和出站交通流时间序列数据，用于实现节假日期间高速公路交通流多步预测。为了方便问题描述，首先定义两个关键参数。

**定义1（收费站进出站交通流时间序列矩阵）：**原始高速公路收费站进出站数据是以60分钟为时间粒度进行统计的，其中包含收费站位置、收费站ID、时间段、进站交通流、出站交通流的信息。以收费站进站车流为例，假设表示收费站在第个时间间隔内的收费站进站交通流量观测值，则高速公路收费站进站交通流时间序列可以被表示为：

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

其中表示整个时段内高速公路收费站的进站车流时间序列，*N*表示高速公路收费站数量，*T*代表整个时段内总时间步数量。同理，高速公路收费站出站交通流时间序列可以表示为下式。

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

**定义2（进出站交通流时间序列增广矩阵）：**为了有效学习节假日期间的进出站流特性，本研究同时使用节假日的历史交通流数据和节前交通流数据以准确预测未来节假日*k*个时间步的交通流状态。具体来说，同一节假日假期的历史交通流数据指的是与当前拟预测交通流相同时刻的去年假期期间交通流数据。设当前时间段为*t*，选取*T*个时间步作为交通流数据的总时间步数量，构建节假日交通流时间序列增广矩阵(以进站流数据为例)，可以定义为表达式(3)。

|  |  |
| --- | --- |
|  | (3) |

其中 表示节假日的历史交通流数据，表示拟预测时间步数量，T表示节前交通流数据的总时间步数量。

**目标方程：**在第 个时间步，高速公路节假日期间收费站进站和出站交通流的增广时间序列矩阵分别为 ，根据该时间序列矩阵，预测未来个时间步的节假日期间进站和出站交通流，该问题可以被定义为公式(4)，其中表示需要训练的多任务节假日高速公路进出站流多部预测模型，和分别为未来时刻的进出站流。

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

2.2 模型构建

为充分捕捉高速公路收费站进出站交通流复杂动态的时空特征，本文构建了时空注意力交互网络（ST-Cross-Attn），主要由进站流分支、出站流分支以及注意力交互分支构成，其网络结构如下图所示。其核心思想是通过进出站交通流的时空信息的交互，增强二者的特征表示，从而有效建模进出站交通流的时空特征。具体地，进站流分支用于处理进站交通流的时间序列数据，通过双向卷积长短时记忆网络（Bi-ConvLSTM）从正向和反向充分捕捉进站流的时空依赖性；类似地，出站流分支用于处理高速公路出站流时间序列数据，通过双向卷积长短时记忆网络充分挖掘出站流的时空依赖性；最后，进站流分支的输出和出站流分支的输出输入到注意力交互分支中，通过注意力交互分支共同学习进站流和出站流的交互信息，准确建模二者间的关联性，增强特征表示。基于构建的时空注意力交互网络，本研究进一步提出了高速公路收费站进出站交通流的预测框架，旨在准确地协同预测未来*k*个时间步的进出站交通流。接下来将详细介绍构成时空注意力网络的进出站流分支、注意力交互分支，最后介绍基于时空注意力交互网络的协同预测框架。



图1 时空注意力交互网络(ST-Cross-Attn)结构图

**(1) 进站流/出站流分支**

为准确预测节假日期间未来 个时间步的收费站进站流量，构建节假日交通流时间序列增广矩阵 ，作为该分支的输入。该增广矩阵充分利用了历史节假日交通流数据，有利于提升模型对节假日特征的学习。

既有研究[18-19]证明了卷积长短时记忆网络（ConvLSTM）[20]在捕捉时序数据的时空特征方面具有较大优势，具体可以表示为以下公式，

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

其中，表示模型输入，表示单元输出，表示隐藏状态，分别表示输入门、遗忘门以及输出门，表示卷积权重，是偏置项，是sigmoid函数，表示双曲正切函数，“\*”表示卷积算子，“”表示哈达玛乘积。

然而非常规时序数据具有随机性和突发性，卷积长短时记忆网络无法有效建模复杂的时空依赖性，导致其在非常规时间序列数据预测方面的性能不佳。进一步分析非常规时间序列数据，对于非常规场景下的正向时间序列而言，其时序数据特征具有一定的连续性；对于非常规场景下的反向时间序列而言，其时序数据具有一定的可追溯性。为了完整挖掘非常规场景下时间序列数据的时空特征，本文提出一个双向卷积长短时记忆网络（Bi-ConvLSTM），其模型示意图具体如下：



图2 双向卷积长短时记忆网络结构图

具体来说，经过预处理的进站流时间序列数据输入到进站流分支后，双向卷积长短时记忆网络分别从正向和反向两个方向对其进行建模，以充分捕捉进站流的时空特征；接着对前向和反向卷积长短时记忆网络捕捉到的时空特征进行拼接，并通过线性全连接操作充分融合。最后，利用ReLu激活函数进一步处理特征矩阵中数据，从而获得进站流的复杂隐藏状态。

与进站流类似，高速公路收费站出站流也同样具有复杂动态的时空依赖性特征。为了捕捉这种复杂的时空依赖性，出站流分支同样采用双向卷积长短时记忆网络模块分别从正向和反向对出站流的时空特征进行建模，最终获得出站流的复杂隐藏状态。另外，使用层标准化操作提升训练速度、提高网络稳定性；使用残差连接消除模型在训练的过程中出现梯度消失或梯度爆炸等情况。

**(2) 注意力交互分支**

现有研究对地铁进出站流间的关系进行深入研究，发现进站客流与出站客流间存在关联[21]。实际上，高速公路收费站的进出站流间也存在一定的因果性和关联性。具体来说，当前时刻某个收费站的进站车流量可能会影响未来时刻其他收费站的出站车流量，本研究将这种关系称为进站流与出站流间的因果性。这种关系在节假日期间主要受出行游玩行为影响，比如人们从某个收费站出发前往旅游景点游玩，从而导致某个收费站站的出站车流骤增。另外，同一个收费站的进站流和出站流存在较强关联性。例如城域通勤中，高速公路收费站的进出站流呈现明显的潮汐特征，早高峰居住区域附近收费站点的进站流远高于其出站流，晚高峰则相反。总的来说，高速公路收费站进站流和出站流之间存在丰富的相互信息，充分利用这些相互信息对收费站进站流和出站流进行协同预测有利于提升预测精度。

为此，在利用双向卷积长短时记忆网络分别获得进出站交通流的隐藏状态后，本文进一步研究二者间的相互关系。具体来说，本研究设计了注意力交互分支，通过利用注意力机制挖掘进站流和出站流的关联，实现进站流隐藏状态和出站流隐藏状态间相互信息的传播，最终有效建模收费站进站流和出站流之间的内在交互过程。注意力机制[10]最早提出用于自然语言序列处理（NLP）任务，其可以在不受节点距离影响的情况下，准确计算两个节点间的相关性，具有较强的相关性特征挖掘能力，因此，本研究对传统注意力机制进行改进，提出交互注意力模块用以联合建模进站流和出站流的时空特征，其模型框架图如下图所示，由两个交互注意力子模块构成，其中右侧的子模块将进站流分支的信息传递到出站流分支；左侧的子模块则将出站流分支信息传递到进站流分支。此处以进站流分支输出的隐藏状态与出站流分支输出的隐藏状态的交互为例，对注意力交互分支的计算流程进行详细介绍。首先，进站流隐藏状态和出站流隐藏状态分别使用卷积嵌入层进行处理从而分别获得查询嵌入（Query）、键嵌入（Key）以及值嵌入（Value），

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

其中，表示卷积核大小为的二维卷积操作，W表示相应的卷积权重。与进出站流的隐藏状态相同，查询嵌入、键嵌入以及值嵌入均具有相同的维度。接着，根据注意力机制，注意力交互分支分别计算进站流的信息传播系数以及出站流的信息传播系数，这些信息传播系数可以动态决定进站流数据与出站流数据间的信息传递情况，

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

其中，Softmax表示激活函数，T表示矩阵的转置。在求得信息传播系数后，进站流与出站流的交互注意力计算如下，

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

在获得进站流与出站流的交互注意力和后，通过元素相加的方式将交互注意力与隐藏状态进行融合，从而获得进站流和出站流的增强隐藏状态，

|  |  |  |
| --- | --- | --- |
|  |  | (9) |

通过进出站流间的信息交互，不仅可以增强进站交通流与出站交通流的隐藏状态，而且有利于捕捉进站流与出站流间的因果性和关联性，进而辅助模型准确预测未来的进出站交通流。



图3 注意力交互分支框架

**(3) 序列到序列框架(Seq2Seq)**

上述提出的双向卷积长短时记忆网络有效地挖掘了高速公路收费站进站流和出站流的复杂时空依赖性，提出的交互注意力机制则深入捕捉进站流与出站流间的因果性和关联性，可以为后续的预测提供有用的交互信息。基于这两个模块，本研究提出一个统一的高速公路收费站进出站流预测框架，旨在实现未来多个时段的进站车流和出站车流的同步预测。

具体来说，该框架是一个典型的序列到序列（Seq2Seq）的预测框架，通过编码-解码过程实现多步预测，其中编码器和解码器均为双向卷积长短时记忆网络，其框架示意图如下图所示。此处以进站流为例，对编码、解码的完整过程进行介绍，出站交通流时间序列的处理过程与进站流完全一致。假设初始时刻为*t*，高速公路所有收费站的进站交通流为，上一时刻进站交通流的隐藏状态为，将和输入到双向卷积长短时记忆网络中编码得到当前时刻的进站交通流隐藏状态，每个时刻的隐藏状态均包含了进站交通流的历史时空特征信息。在编码得到进站交通流的隐藏状态编码与出站交通流的隐藏状态编码后，将二者输入到注意力交互分支中，通过交互注意力机制学习它们之间的因果性与关联性，从而利用丰富的相互信息，获得增强的隐藏状态编码、。进一步地，将获得的增强隐藏状态编码作为历史隐藏状态，与下一时刻的进站交通流一同输入到双向卷积长短时记忆网络中继续编码获得下个时刻的隐藏状态，并通过注意力交互分支获得下一时刻的增强隐藏状态编码，依次迭代直到获得当前时刻的增强隐藏状态编码。需要注意的是，在初始时刻，双向卷积长短时记忆网络输入的进站交通流的隐藏状态为0。上述过程为进站交通流的特征编码过程，出站交通流的特征编码过程与其完全一致。

在获得当前时刻增强的进站交通流隐藏状态编码后，利用解码器对其进行解码，并通过全连接层对解码器的输出进行映射操作以获得未来时刻的特征表示，最后与历史假期相同时刻的进站交通流进行残差连接，获得未来下一时刻的进站交通流预测值。进一步地，将*t+1*时刻的预测值和隐藏状态输入到解码器中获得下一时刻的隐藏状态，并依次通过全连接和残差连接获得*t+2*时刻的进站交通流预测值。根据上述解码过程，反复迭代T次最终获得未来T个时间步的进站交通流预测值。类似地，出站交通流也可以通过上述迭代过程生成未来T个时间步的出站交通流预测值。



图4 序列到序列（Seq2Seq）预测框架

**3 模型实验**

3.1 数据集描述

本论文采用2021年（2022年受疫情影响数据量存在较大偏差不采用）、2023年和2024年劳动节期间的山东省全省高速公路收费站进出站流数据集。该数据集一共包含382个高速公路收费站点在劳动节前后5天的数据，即2021年4月26日-2021年5月5日、2023年4月26日-2023年5月5日以及2024年4月26日-2024年5月5日。每条数据中均记录了收费站名称、节点ID、日期、小时、进站流以及出站流信息。由于存在不可能避免的数据收集误差，在提取进出站流时间序列数据之前，需要对原始数据进行数据清洗，例如一些缺失站点信息的进出站流数据或者较长时间内进出站流为零的数据。在完成数据清洗后，依据章节1.1问题描述部分对于时间序列数据的定义，以60分钟为时间粒度，分别提取收费站进站交通流和出站交通流的时间序列数据。

本文旨在研究节假日期间高速公路收费站进站、出站交通流多步预测问题。为了更好捕捉高速公路交通流节假日特征，挖掘节假日对高速公路交通流的影响，本研究提出进出站交通流时间序列增广矩阵作为模型的输入。以23年假期数据预测为例，假设当前时刻为*t*，需要预测5月1日的交通流数据，那么当前时间序列增广矩阵由21年5月1日的交通流数据以及当前时间刻*t*之前的T个时间步（本实验T设置为5\*24）的交通流数据组成。通过使用历史节假日数据，模型可以更好地捕捉假期交通流时空特征，从而提升模型的预测效果。

3.2 模型配置

本研究提出的模型基于Pytorch实现，进站交通流和出站交通流数据按照5：3：2的比例划分训练集、验证集以及测试集。通过多次试验后选择批次大小为32，学习率为0.001，优化器为Adam。训练的过程中，选取均方误差(MSE)作为损失函数。由于本研究采取多任务学习，因此模型的损失函数由两部分加权求和获得，分别是进站交通流的均方误差以及出站交通流的均方误差，进站交通流与出站交通流的权重在该损失函数中保持一致，具体见公式(17)：

|  |  |  |
| --- | --- | --- |
|  |  | (10) |

其中，和分别表示进站交通流和出站交通流的样本量，和分别表示进出站交通流的真实值，和分别表示预测的进出站交通流。另外，本文选取均方根误差(RMSE)，平均绝对误差(MAE)以及加权平均绝对百分比误差 (WMAPE)作为模型预测效果的评估指标，各指标计算公式如(11)-(13)。为方便书写，此处将进出站交通流的真实值统一使用表示，将进出站交通流的预测值统一使用表示，进出站交通流的样本数量统一使用表示。

|  |  |  |
| --- | --- | --- |
|  |  | (11) |
|  |  | (12) |
|  |  | (13) |

3.3 基准实验设置

(1) SVR (Support Vector Regression): 作为经典的机器学习模型，支持向量回归模型被广泛应用于交通预测领域[22]，主要利用支持向量机技术对时间序列数据进行回归分析。本实验使用内核函数‘rbf’，设置误差范围参数 epsilon=0.005，正则化参数 C=3。

(2) XGBoost (eXtreme Gradient Boosting)[23]: 梯度提升决策树算法是一种高效的机器学习算法，将多个弱学习器通过一定的方法集合为一个强学习器，从而提升整体效果，近年来在交通预测领域被广泛应用。使用‘gbtree’作为提升器，设置树的最大深度 max\_depth=6，学习率为0.1。

(3) BPNN (Back Propagation Neural Network): 反向传播神经网络作为最经典的神经网络之一，被证明可以有效捕捉交通流的非线性特征[24]。本研究搭建的BPNN由两层全连接层组成，每层分别有1024和512个神经元。模型的优化器是Adam，学习率为0.0005。

(4) CNN (Convolutional Neural Network): 卷积神经网络在捕捉数据的空间相关性方面具有良好的性能[25]，卷积核能够有效捕捉时间序列特征之间的空间依赖性。本研究首先利用卷积核大小为3×3的卷积神经网络层建模空间特征，随后使用全连接层进行预测，并采用ReLU作为激活函数。

(5) LSTM (Long Short-Term Memory): 长短时记忆网络可以有效解决常规循环神经网络中长时依赖问题，能有效提取到时间序列中更长时间的变化趋势，因此在交通客流预测领域具备一定优越性[26]。该模型设置一层LSTM层，其隐藏层层数为4；全连接层层数为1，并采用ReLU函数激活。

(6) GRU (Gated Recurrent Unit): 门控循环单元是一种经典的循环神经网络，可以有效解决长时依赖问题以及反向传播中梯度消失问题，被广泛应用于时间序列处理问题中[27]。相比于LSTM，GRU具有更加简洁的结构，在提升训练速度的同时保持较好的效果。该模型由3个隐藏层和1个全连接层组成。每个隐藏层包含128个神经单元。

(7) ST-ResNet (Spatial Temporal Residual Network): 该网络采用残差卷积单元建模交通流的时空相关性，并使用三个分支对交通流的近期、周期以及趋势时间特征进行学习。本研究利用一层残差卷积层学习时空特征，卷积核大小为3×3，并使用全连接层预测未来交通流。

(8) Transformer: 该模型[10]最早于2017年提出用于处理自然语言(NLP)，其可以有效捕捉任意节点间的相关性而不用考虑节点间的距离。因此Transformer近年来被广泛用于时间序列处理。本研究搭建了包含2层encoder和2层decoder的Transformer框架。其中，多头数设置为4，特征大小d\_model为256。Transformer的输出将输入到全连接层进行预测。

**4 预测结果分析**

本节将从模型的网络级预测效果、车站级预测[28]效果以及消融实验等几个方面对本文提出的模型进行全面分析。

4.1 总体预测分析

基于前文所述节假日期间高速公路收费站进站流和出站流数据集，所有模型均采用固定输入步长（历史5天，120个时间步）的流量数据和历史节假日交通流量数据（历史步长与预测步长保持一致）作为输入，分别预测未来1天（24个时间步）、2天（48个时间步）和3天（72个时间步）的流量数据。基准模型均采用独立且相同的模块分别预测收费站未来进站流量和出站流量。预测效果如表1和表2所示。

首先，基于机器学习的SVR和XGBoost模型在不同预测时间步数下的预测性能均低于其他模型。由于这些模型进行多步预测时，每个预测时间步均对应一个独立的模型，无法直接处理时间序列数据的时间依赖性，表明不适合高速公路节假日多步进出站流量预测。

其次，虽然BPNN能够捕捉节假日流量的非线性关系，但由于缺乏对时空依赖性的建模，导致其表现仅优于基于机器学习的模型。LSTM、GRU和Transformer通过捕捉时间依赖性，实现了比BPNN更好的预测性能。Transformer利用自注意力机制捕捉长时间依赖，表现优于LSTM和GRU。CNN通过卷积操作捕捉节假日流量的局部站点间空间关系和时间依赖，但随着时间步数的增加，其预测性能下降，因为CNN难以捕捉全局流量信息。ST-ResNet通过全局建模时间和空间依赖性，预测精度优于其他模型，但未能考虑进出站流量之间的交互关系。

最后，为了捕捉节假日期间复杂的高速公路流量模式，本文提出的模型采用了Bi-ConvLSTM同时捕捉空间依赖性和全局时间依赖性。同时，进站流量和出站流量具有显著的相关性，本文通过Cross-attention机制捕捉节假日高速公路进站流量与出站流量之间的关联，实现了进出站流量特征的共享。结果表明，与基准模型相比，模型在不同预测时间步上表现出最佳的预测性能，证明了本文所提出模型的预测准确性和鲁棒性。

表 1进站流预测结果(Inflow Dataset)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **进站流** | | | | | | | | | |
| 模型 | 1天 (24个时间步) | | | 2天 (48个时间步) | | | 3天 (72个时间步) | | |
| RMSE | MAE | WMAPE | RMSE | MAE | WMAPE | RMSE | MAE | WMAPE |
| SVR | 205.53 | 129.47 | 41.93% | 204.09 | 128.69 | 41.31% | 201.46 | 127.33 | 40.71% |
| XGBoost | 184.70 | 100.13 | 39.28% | 184.64 | 101.93 | 38.56% | 184.44 | 102.43 | 42.74% |
| BPNN | 161.26 | 85.82 | 33.06% | 191.70 | 103.15 | 39.35% | 187.66 | 96.23 | 34.56% |
| CNN | 158.84 | 84.72 | 30.77% | 166.27 | 87.72 | 31.51% | 184.91 | 99.41 | 35.73% |
| LSTM | 163.29 | 84.88 | 30.66% | 159.66 | 83.48 | 29.82% | 161.41 | 81.69 | 29.16% |
| GRU | 180.55 | 94.12 | 34.06% | 175.07 | 90.44 | 32.36% | 159.82 | 84.30 | 30.14% |
| ST-ResNet | 151.07 | 79.28 | 28.78% | 161.18 | 85.18 | 30.58% | 180.75 | 99.57 | 35.76% |
| Transformer | 145.81 | 87.24 | 31.68% | 148.63 | 86.34 | 30.98% | 153.32 | 80.37 | 28.86% |
| ***Our Model*** | ***136.89*** | ***66.91*** | ***28.55%*** | ***142.32*** | ***70.27*** | ***28.47%*** | ***148.34*** | ***72.08*** | ***27.47%*** |

表2 出站流预测结果(Outflow Dataset)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **出站流** | | | | | | | | | |
| 模型 | 1天 (24个时间步) | | | 2天 (48个时间步) | | | 3天 (72个时间步) | | |
| RMSE | MAE | WMAPE | RMSE | MAE | WMAPE | RMSE | MAE | WMAPE |
| SVR | 177.89 | 114.13 | 37.42% | 232.94 | 150.77 | 44.32% | 175.65 | 111.58 | 36.16% |
| XGBoost | 161.32 | 89.46 | 37.51% | 194.20 | 110.49 | 40.12% | 164.31 | 91.28 | 37.76% |
| BPNN | 142.34 | 75.65 | 28.41% | 158.47 | 86.28 | 33.53% | 152.34 | 81.43 | 29.87% |
| CNN | 126.68 | 69.58 | 25.69% | 145.30 | 81.07 | 29.74% | 161.56 | 94.59 | 34.69% |
| LSTM | 133.85 | 69.04 | 25.49% | 132.42 | 69.41 | 25.45% | 139.51 | 72.53 | 26.61% |
| GRU | 140.69 | 73.59 | 27.18% | 136.94 | 72.15 | 26.47% | 130.81 | 73.20 | 26.85% |
| ST-ResNet | 132.01 | 71.03 | 26.23% | 146.56 | 82.07 | 30.10% | 162.27 | 92.54 | 33.94% |
| Transformer | 117.49 | 75.56 | 27.89% | 127.39 | 70.71 | 25.94.% | 132.44 | 76.64 | 27.38% |
| ***Our Model*** | ***116.63*** | ***54.70*** | ***24.28%*** | ***122.86*** | ***59.54*** | ***25.37%*** | ***127.39*** | ***61.99*** | ***24.76%*** |

4.2 车站级预测效果分析

为验证模型在单个收费站层面的具体预测效果，本文选取三个典型收费站，该三个收费站在不同预测时间步数量下预测流量与真实流量的对比如图5-图7所示。

第一个收费站为临沂市蒙阴收费站，节假日期间该站的进站流量和出站流量相当，为均衡性收费站。如图5所示，该收费站的进出站流量在工作日具有明显的早晚高峰特征，节假日进出站流量显著高于工作日流量。在不同预测时间步下，预测值均与真实值拟合效果较好，表明模型具有良好预测性能。

第二个收费站为济南市埠村收费站，节假日期间该站的进站流量大于出站流量，为流入型收费站。如图6所示，该收费站的进站和出站流量在工作日具有显著的规律性，表现出明显的早晚高峰特征。节假日期间，进站流量略高于工作日，而出站流量与工作日相当，节假日期间的流量增长并不显著。在所有预测时间步长下，预测曲线与真实值高度吻合，预测效果良好。

第三个泰安市彭集收费站，节假日期间该站出站流量大于进站流量，为流出型收费站。如图7所示，该站的工作日进出站流量明显低于前两个收费站，但在节假日期间，进出站流量增长幅度显著。尽管如此，本文所提出的模型可以准确捕捉交通流变化趋势，具有较强的鲁棒性。

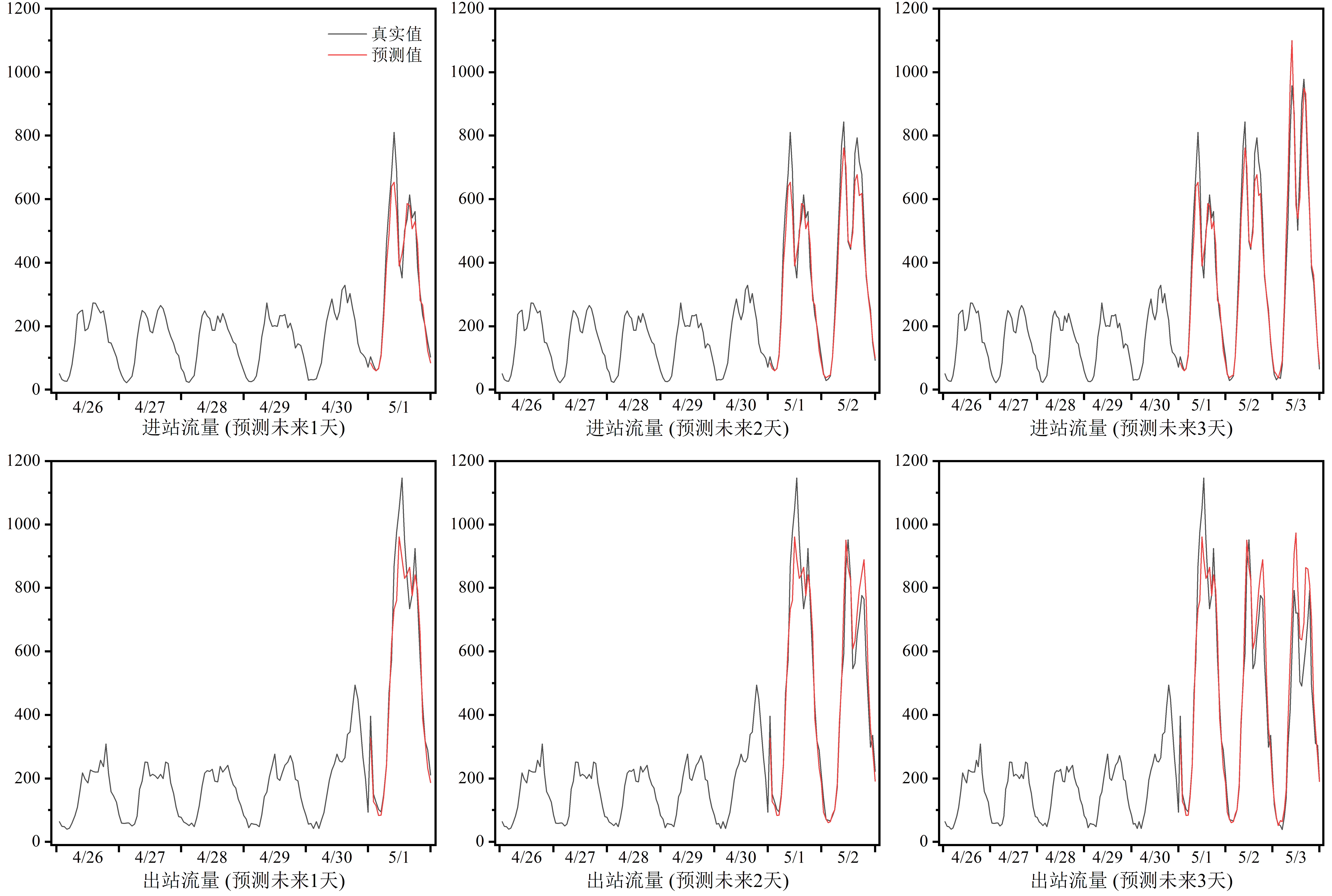


图5 蒙阴收费站流量预测值-真实值对比

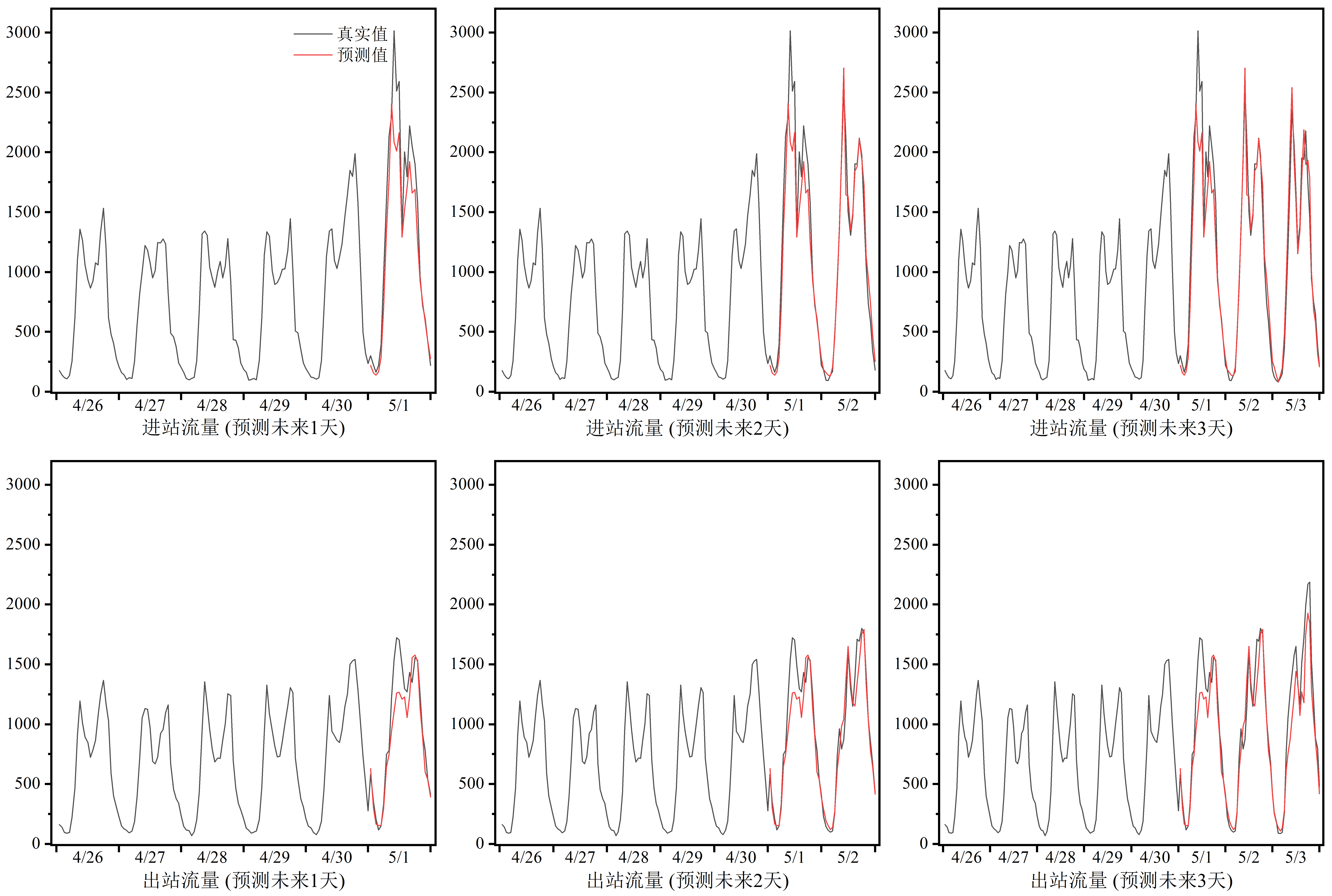


图6 埠村收费站客流预测值-真实值对比

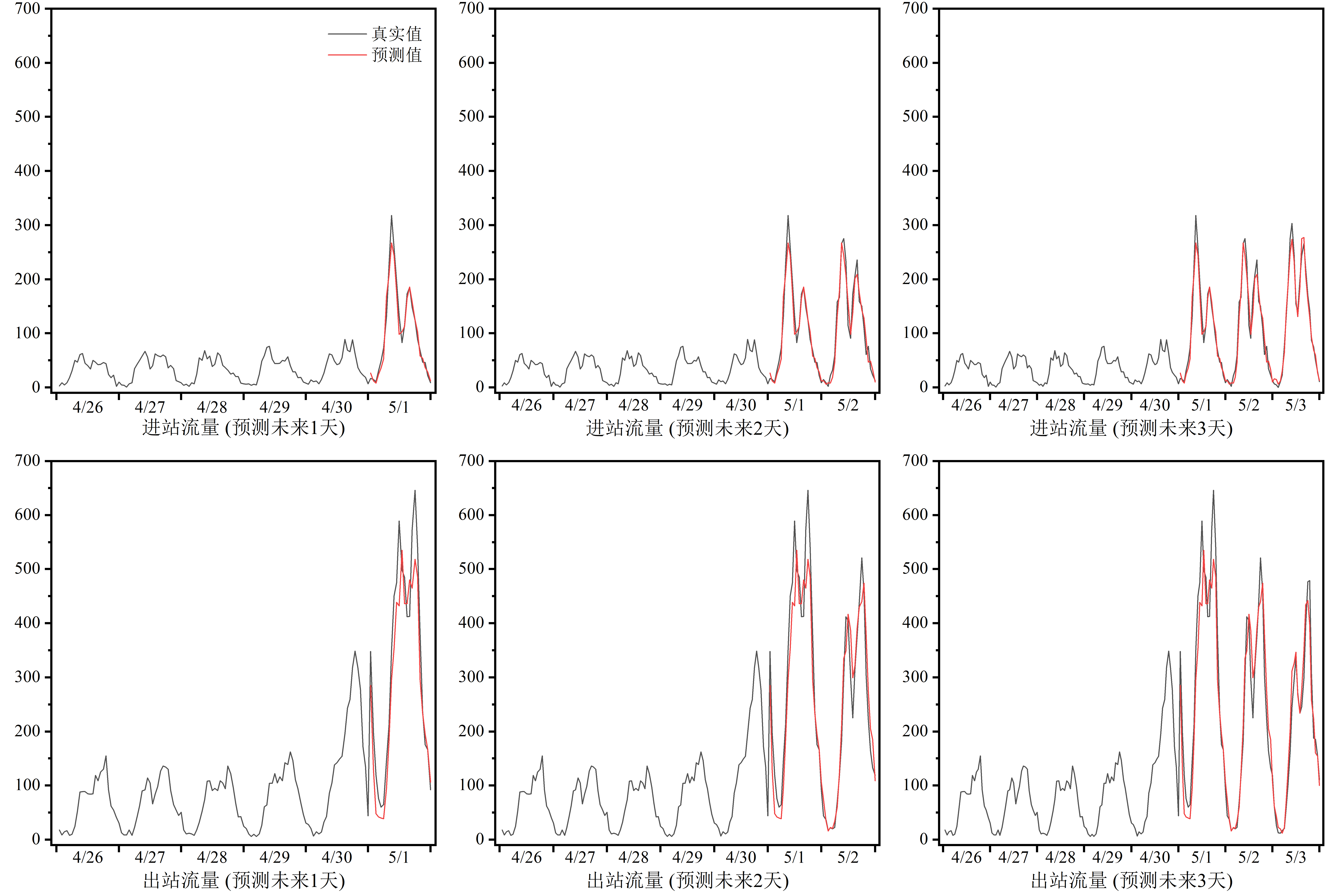


图7 彭集收费站流量预测值-真实值对比

4.3 消融实验

为了验证本文所提出的时空注意力交互模型中各个模块的有效性与优越性，本节进行了一系列消融实验，具体变体如下所示。

**变体1: ST-Cross-Attn-No Bi：**使用卷积长短时记忆网络模块替换时空注意力交互模型中的双向卷积长短时记忆网络模块，其余配置基本不变，探究双向建模方式对于特征挖掘的影响。

**变体2: ST-Cross-Attn-LSTM：**使用双向长短时记忆网络（Bi-LSTM）模块替换时空注意力交互模型中的双向卷积长短时记忆网络模块，其余配置基本不变，探究卷积运算在捕捉时空特性方面的能力。

**变体3: ST-Cross-Attn-No Attn：**在所提出模型中，删除注意力交互模块，探究交互注意力在提升特征表示方面的能力。需要注意的是，在去除该模块后，进站交通流和出站交通流的隐藏特征不再进行交互，此时模型可以视为两个单任务学习模型的简单合，由于模型的损失函数中仍然包含进站交通流部分和出站交通流部分，因此与单任务学习还是存在区别。

**变体4: ST-Cross-Attn-No holiday：**为学习节假日特征，本研究提出了进出站交通流增广时间序列矩阵，利用历史假期交通流辅助预测未来假期的交通流。此处使用原始的进出站交通流时间序列矩阵对未来交通流进行预测，旨在探究历史节假日交通流数据对与当前节假日交通流预测效果的提升。

表 3 消融实验进站交通流预测结果精度

Table 3 Inflow Prediction Result Evaluation of Variants

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **进站流** | | | | | | | | | |
| 变体 | 1天 (24个时间步) | | | 2天 (48个时间步) | | | 3天 (72个时间步) | | |
| RMSE | MAE | WMAPE | RMSE | MAE | WMAPE | RMSE | MAE | WMAPE |
| 变体1(No Bi) | 137.55 | 68.56 | 29.25% | 143.21 | 72.04 | 29.19% | 155.94 | 87.61 | 33.42% |
| 变体2(LSTM) | 152.92 | 83.85 | 35.87% | 158.29 | 86.19 | 34.96% | 160.47 | 85.26 | 32.53% |
| 变体3(No Attn) | 139.84 | 74.35 | 31.71% | 145.76 | 77.67 | 31.47% | 169.76 | 83.35 | 31.80% |
| 变体4(No holidays) | 165.83 | 76.84 | 32.10% | 160.47 | 75.03 | 32.08% | 153.39 | 74.12 | 30.01% |
| ***ST-Cross-Attn*** | ***136.89*** | ***66.91*** | ***28.55%*** | ***142.32*** | ***70.27*** | ***28.47%*** | ***148.34*** | ***72.08*** | ***27.47%*** |

表 4 消融实验出站交通流预测结果精度

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **出站流** | | | | | | | | | |
| 变体 | 1天 (24个时间步) | | | 2天 (48个时间步) | | | 3天 (72个时间步) | | |
| RMSE | MAE | WMAPE | RMSE | MAE | WMAPE | RMSE | MAE | WMAPE |
| 变体1(No Bi) | 117.23 | 55.88 | 24.79% | 124.22 | 61.72 | 26.87% | 132.51 | 72.22 | 28.82% |
| 变体2(LSTM) | 135.35 | 71.09 | 31.55% | 139.64 | 73.63 | 31.36% | 142.27 | 73.24 | 29.24% |
| 变体3(No Attn) | 119.10 | 60.01 | 26.62% | 125.71 | 64.83 | 27.61% | 162.37 | 79.87 | 31. 89% |
| 变体4(No holidays) | 167.52 | 65.37 | 29.02% | 154.16 | 72.71 | 30.83% | 138.08 | 71.69 | 28.63% |
| ***ST-Cross-Attn*** | ***116.63*** | ***54.70*** | ***24.28%*** | ***122.86*** | ***59.54*** | ***25.37%*** | ***127.39*** | ***61.99*** | ***24.76%*** |

表3和表4记录了所有消融实验结果，根据上述实验结果，得出以下的实验结论：

(1) **变体1将双向卷积长短时记忆网络替换成单向卷积长短时记忆网络，仅从单一方向挖掘交通流的时空依赖性。**由预测结果分析可知，模型的预测精度出现下降，这是因为单向建模无法充分挖掘时序数据的时间关联性和因果性，因此证明了双向建模方式的有效性。同时，变体1随着预测时间步的增加，预测误差逐渐变大，展示了双向建模在多步预测上的优越性，表明在面对较长预测时间步长的情况下，双向建模可以充分挖掘长时的时空依赖性，避免时间步数过长导致特征学习不足。

(2) **变体2将双向长短时记忆网络替换成双向长短时记忆网络，仅挖掘交通流时间序列数据的时间依赖性。**由预测结果分析可知，模型的预测精度明显下降，这表明使用LSTM进行特征挖掘时，模型只能捕捉有限的时间特征，无法充分建模收费站进出站交通流的空间特征。事实上，收费站进出站交通流间具有复杂的空间特征，例如市域通勤收费站点间存在显著的相关性，这些站点的进站交通流与出站交通流之间的潮汐特征明显。卷积长短时记忆网络可以有效利用卷积操作捕捉到目标站点的空间特征，从而有效提升预测精度。

(3) **变体3删除了原模型的注意力交互机制，不考虑进站流和出站流之间的关联性。**由实验结果分析可知，该变体在不同预测步长下的预测精度均出现下降，且随着预测步长的增加，预测精度下降程度增大。这是因为注意力机制在捕捉目标节点与其他节点的相关性时可以忽略节点间的距离长度，从而有效计算长距离外节点的相关性。在删除交互注意力机制后，模型计算目标节点与距离较远节点间相关性的能力减弱，随着预测步长增加，无法有效计算目标节点与更远的时间序列节点间的相关性，导致预测精度大幅下降。同时，交互注意力机制可以共享进站交通流与出站交通流的时空信息，增强特征表示，在删除该模块后，进站交通流与出站交通流的特征信息不再共享，无法实现时空特征交互，使预测精度下降。因此，交互注意力机制不仅有利于计算不同时间序列节点间的相关性，还可以共享进出站流特征信息，增强特征表示，从而提升预测精度。

(4) **变体4的输入由进出站交通流的增广时间序列矩阵替换为了进出站交通流时间序列矩阵，不考虑历史节假日的进出站交通流时间序列信息。**由实验结果分析可知，当模型不考虑历史节假日进出站交通流信息后，预测精度出现不同程度的下降，且当预测时间步为24时，预测效果最差，随着预测步长增加，预测效果提升。分析原因，主要是因为节假日特征通常出现在假期的第一天，当天交通流往往骤增，而后交通流相较日常的增幅变小，特征不如第一天明显。在不考虑历史节假日特征的情况下，预测时间步为24恰好为假期第一天，模型难以学习假期特征，所以第一天的预测效果最差；随着预测时间步增加，后两天的假期特征不如第一天明显，所以模型在后面两天的预测精度逐步提升。因此，出现了上述的实验结果，也进一步证明了本文构建的增广时间序列矩阵的有效性。

最后，与表格1对比可以看出，虽然消融实验中提出的变体模型的预测精度有所下降，但预测效果优于大部分基准模型，表明模型中的每个模块都是不可或缺的，可以有效提升模型的预测性能。

4.4 多任务学习效果分析

本研究采用多任务学习方式训练模型，通过共享进出站交通流的时空特征信息，增强二者的关联性，并且充分捕捉进出站流时空特征，从而实现精确的进出站交通流协同预测。为了验证多任务学习方式的有效性，本研究进一步分析多任务学习方式与单任务学习方式下模型的预测精度。具体来说，除了使用本文提出的基于多任务学习方式的预测模型协同预测进出站交通流，还使用单任务学习方式分别对进站交通流和出站交通流进行预测，比较两种不同预测方式下的预测精度。单任务学习方式下，模型中的注意力交互模块使用原始的注意力机制替换，即进出站流通过双向卷积长短时记忆网络挖掘时空依赖性后，使用原始的注意力机制进行处理，增强特征表示。经过实验，最终的预测结果如下表所示。

表5 多任务学习与单任务学习进出站交通流预测结果精度

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1天 (24个时间步) | | | 2 天(48个时间步) | | | 3天(72个时间步) | | |
|  | RMSE | MAE | WMAPE | RMSE | MAE | WMAPE | RMSE | MAE | WMAPE |
| 多任务(进站流) | *136.89* | *66.91* | *28.55%* | *142.32* | *70.27* | *28.47%* | *148.34* | *72.08* | *27.47%* |
| 单任务(进站流) | 138.44 | 71.10 | 30.32% | 144.04 | 73.90 | 29.94% | 149.66 | 75.03 | 28.59% |
| 多任务 (出站流) | *116.63* | *54.70* | *24.28%* | *122.86* | *59.54* | *25.37%* | *127.39* | *61.99* | *24.76%* |
| 单任务(出站流) | 139.93 | 74.74 | 31.87% | 145.42 | 76.95 | 31.18% | 152.67 | 80.72 | 30.78% |

由上述预测结果可知，在不同预测步长的情况下，基于多任务学习方式的进出站交通流协同预测模型的预测效果均优于单任务学习方式的进出站交通流预测效果，这表明多任务学习通过进出站交通流的时空特征信息共享交互，学习二者间的关联性，增强了二者的特征表示，从而有效提升了模型的协同预测效果。进一步分析，出站流在多任务学习和单任务学习方式下的预测效果相差较大，这表明进站交通流能够为出站交通流的预测提供更为丰富的时空分布信息，从而使模型更准确地预测出站交通流。

**5 结论**

节假日期间的高速公路交通流预测对于高速公路的运营管理具有重要意义，因此本文以节假日期间高速公路收费站进出站流为研究对象，提出基于多任务学习框架的进出站交通流协同预测模型。具体来说， 该模型以进出站交通流增广时间序列矩阵为输入，充分考虑历史假期的交通流特征，通过基于双向卷积长短时记忆网络和注意力交互模块的时空注意力交互模块充分捕捉进出站交通流的时空依赖性，最终利用“序列到序列”框架实现未来多个时间步的进站交通流和出站交通流的协同预测。本文在五一期间山东省高速公路收费站进出站交通流数据集上进行实验，将本文提出的模型与其他八个先进的基线模型进行比较，实验结果表明模型具有良好的预测性能和鲁棒性，可以准确地预测假期期间高速公路的交通流；另外，大量的消融实验证明模型中每个模块的重要性以及多任务学习框架的有效性。

然而本研究也存在部分不足，例如在研究假期期间的交通流预测时，模型仅仅考虑了历史假期交通流数据辅助预测，例如社交媒体上相关推文、旅游景点热度等外界因素对会对假期交通流产生影响，后续研究可以考虑相关因素以提升预测精度。

**参考文献：**

[1] 田俊山,曾俊铖,丁峰,等.基于时空关系的高速公路交通流量预测[J].工程科学学报,2024,46(09):1623-1629.DOI:10.13374/j.issn2095-9389.2023.10.24.004.

[2] 冯腾.基于EMD和GS-SVM融合的节假日高速公路交通流预测研究[D].长安大学,2016.

[3] Ma Z L, Xing J P, Mesbah M, et al. Predicting Short-term Bus Passenger Demand using a Pattern Hybrid Approach[J]. Transportation Research Part C: Emerging Technologies, 2014, 39: 148-163.

[4] Zhang J B, Zheng Y, Qi D K et al. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction[C]. Thirty-first AAAI conference on artificial intelligence, 2017, 31(1).

[5] Li Y G, Yu R, Shahabi C et al. Diffusion Convolutional Recurrent Neural Network: Data-driven Traffic Forecasting[C]. International Conference on Learning Representations (ICLR), 2018.

[6] 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.

[7] Zhang J L, Chen F, Guo Y N, et al. Multi-Graph Convolutional Network for Short-Term Passenger Flow Forecasting in Urban Rail Transit[J]. IET Intelligent Transport Systems, 2020, 14: 1210-1217.

[8] Wu Z H, Pan S R, Long G D, et al. Graph WaveNet for Deep Spatial-Temporal Graph Modeling[C]. Proceedings of the 28th International Joint Conference on Artificial Intelligence, 2019, 1907-1913.

[9] Zhang S X, Zhang J L, Yang L X, et al. COV-STFormer for Short-Term Passenger Flow Prediction During COVID-19 in Urban Rail Transit Systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(5): 3793-3811.

[10] Vaswani A, Shazeer N, Parmar N, et al. Attention is All You Need[C]. Advances in Neural Information Processing Systems, 2017.

[11] Jiang J W, Han C K, Zhao W X, et al. Pdformer: Propagation Delay-aware Dynamic Long-range Transformer for Traffic Flow Prediction[C]. Proceedings of the AAAI conference on artificial intelligence, 2023, 37(4): 4365-4373.

[12] Reza S, Ferreira M C, Machado J, et al. A Multi-head Attention-based Transformer Model for Traffic Flow Forecasting with a Comparative Analysis to Recurrent Neural Networks[J]. Expert Systems with Applications 2022, 202: 117275.

[13] 唐郑熠,黄嘉欢,王金水,等.基于自适应扩散图卷积注意力网络的地铁客流预测[J/OL].铁道科学与工程学报,1-13[2024-09-01].https://doi.org/10.19713/j.cnki.43-1423/u.T20240295.

[14] Yan H Y, Ma X L, Pu Z Y, et al. Learning Dynamic and Hierarchical Traffic Spatiotemporal Features with Transformer[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(11): 22386-22399.

[15] Xu M X, Dai W R, Liu C M, et al. Spatial-Temporal Transformer Networks for Traffic Flow Forecasting[J]. arXiv preprint arXiv:2001.02908, 2020.

[16] Zhang S X, Zhang J L, Yang L X, et al. Physics Guided Deep Learning-based Model for Short-term Origin-Destination Demand Prediction in Urban Rail Transit Systems Under Pandemic[J]. Engineering, 2024, 41: 276-296.

[17] Zhang S X, Zhang J L, Yang L X, et al. Spatiotemporal Attention Fusion Network for Short-Term Passenger Flow Prediction on New Year’s Day Holiday in Urban Rail Transit System[J]. IEEE Intelligent Transportation Systems Magazine, 2023, 15(5): 59-77.

[18] Azad R, Asadi-Aghbolaghi M, Fathy M, et al. Bi-directional ConvLSTM U-Net with Densley Connected Convolutions[C]. Proceedings of the IEEE/CVF international conference on computer vision workshops, 2019.

[19] Lin Z H, Li M M, Zheng Z B, et al. Self-Attention ConvLstm for Spatiotemporal Prediction[C]. Proceedings of the AAAI conference on artificial intelligence, 2020, 34(7).

[20] Shi X J, Chen Z R, Wang H, et al. Convolutional LSTM network: A Machine Learning Approach for Precipitation Nowcasting[C], Advances in neural information processing systems, 2015, 18.

[21] Liu L B, Zhu Y Y, Li G B, et al. Online Metro Origin-Destination Prediction via Heterogeneous Information Aggregation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(3): 3574-3589.

[22] Sapankevych N, Sankar R. Time Series Prediction using Support Vector Machines: a Survey[J]. IEEE computational intelligence magazine, 2009, 4(2): 24-38.

[23] Zhang Y, Shi X P, Zhang S, et al. A XGBoost-based Lane Change Prediction on Time Series Data using Feature Engineering for Autopilot Vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10): 19187-19200.

[24] Wang L, Zeng Y, Chen T. Back Propagation Neural Network with Adaptive Differential Evolution Algorithm for Time Series Forecasting[J]. Expert Systems with Applications, 2015, 42(2): 855-863.

[25] Wang K, Li K L, Zhou L Q, et al. Multiple Convolutional Neural Networks for Multivariate Time Series Prediction[J]. Neurocomputing, 2019, 360: 107-119.

[26] Siami-Namini S, Tavakoli N, Namin A S, et al. The Performance of LSTM and BiLSTM in Forecasting Time Series[C]. 2019 IEEE International conference on big data (Big Data), 2019.

[27] Yamak P, Li Y J, Gadosey P, et al. A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting[C]. Proceedings of the 2019 2nd international conference on algorithms, computing and artificial intelligence, 2019.

[28] 张金雷, 陈奕洁, Panchamy Krishnakumari,等.基于注意力机制的城市轨道交通网络级多步短时客流时空综合预测模型[J].地球信息科学学报,2023,25(04):698-713.

**Multi-Task Learning-based Holiday Highway Traffic Flow Prediction**

Chai Qinggang1, Zhang Shuxin2\*, Qiu Hao2, Lv Mengqi3, Zhang Yunfan4, Wang Chengcheng3, Zhang Jinlei2

(1. Shandong Expressway Infrastructure Construction Co., Ltd., Jinan 250098, China; 2. School of Systems Science, Beijing Jiaotong University, Beijing 100044, China; 3. Innovation Research Institute of Shandong Hi-Speed Group Co., Ltd., Jinan 250098, China; 4 Shandong Provincial Communications Planning and Design Institute Group Co., Ltd., Jinan 250098, Chin)

Abstract: Accurate traffic flow prediction can provide critical data support for highway operational decision-making, particularly during holidays when surges in traffic flow pose significant challenges to highway management. To address this issue, this study focuses on highway traffic flow during holiday periods and proposes a spatiotemporal attention interaction network. The network is constructed based on a bidirectional convolutional long short-term memory (BiConvLSTM) network and an attention interaction module, aiming to jointly predict the inbound and outbound traffic flow at highway toll stations during holidays. The model's performance was evaluated on a dataset of toll station traffic flow collected during the May Day holiday in Shandong Province, and its predictive capabilities were compared against eight advanced baseline models across different prediction horizons. For inbound traffic flow prediction, the proposed model outperformed the second-best model by an average of 4.34%, while for outbound traffic flow prediction, it achieved an average improvement of 2.3% over the second-best model. These results demonstrate the model’s excellent predictive performance and its capability to accurately forecast multi-step traffic flow during holiday periods.

Keywords: Road Transportation; Traffic Flow Prediction; Deep Learning; Highways; Traffic Big Data; Spatial-Temporal Features Modeling

1. 基金项目：国家自然科学基金青年基金项目/Young Scientists Fund of the National Natural Science Foundation of China (72201029)。

   作者简介：柴庆刚（1970-），男，山东人，高级工程师，研究方向为高速公路建设与管理。

   \*通讯作者：章树鑫（1999-），男，广东人，博士，研究方向为人工智能与交通大数据挖掘及短时客流预测，[22110249@bjtu.edu.cn](mailto:22110249@bjtu.edu.cn)。

   收稿日期：XXXX－XX－XX [↑](#footnote-ref-1)