

Article

Multi-Task Learning-Based Traffic Flow Prediction Through Highway Toll Stations During Holidays

Xiaowei Liu ¹, Yunfan Zhang ¹, Zhongyi Han ¹, Hao Qiu ², Shuxin Zhang ^{2,*} and Jinlei Zhang ^{2,3,*}

- ¹ Shandong Provincial Communications Planning and Design Institute Group Co., Ltd., Jinan 250101, China
² School of Systems Science, Beijing Jiaotong University, Beijing 100044, China
³ National Engineering Research Center of System Technology for High-Speed Railway and Urban Rail Transit, Beijing 100081, China
* Correspondence: 22110249@bjtu.edu.cn (S.Z.); zhangjinlei@bjtu.edu.cn (J.Z.)

Abstract

Accurate traffic flow prediction is essential for highway operations, especially during holidays when surging traffic poses significant challenges. This study focuses on holiday traffic and introduces a spatiotemporal cross-attention network (ST-Cross-Attn) that combines a bidirectional convolutional LSTM (Bi-ConvLSTM) with a cross-attention module to jointly predict toll station inbound flow and outbound flow. Under the multi-task learning framework, the model shares spatial-temporal features between inbound flow and outbound flow, enhancing their representations and improving multi-step prediction accuracy. Using three years of highway traffic flow data during Labor Day from Shandong, China, ST-Cross-Attn outperformed eight state-of-the-art benchmarks, achieving an average improvement of 4.34% in inbound flow prediction and 2.3% in outbound flow prediction. Extensive ablation studies further confirmed the effectiveness of the model's components and multi-task learning framework, demonstrating its potential for reliable holiday traffic forecasting.

Keywords: highway traffic flow prediction; multi-task learning; spatial-temporal features modeling



Academic Editor: Pedro Antonio Gutiérrez

Received: 21 April 2025

Revised: 23 June 2025

Accepted: 1 July 2025

Published: 4 July 2025

Citation: Liu, X.; Zhang, Y.; Han, Z.; Qiu, H.; Zhang, S.; Zhang, J.

Multi-Task Learning-Based Traffic Flow Prediction Through Highway Toll Stations During Holidays. *Technologies* **2025**, *13*, 287.

<https://doi.org/10.3390/technologies13070287>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Highways serve as the backbone of the transportation network, playing a vital role in the integrated transport system [1]. With rapid economic and social development, travel demand has steadily increased, leading to greater highway traffic flow, especially during specific holidays (e.g., Labor Day and National Day). During these periods, toll-free policies often cause a sharp rise in traffic, making widespread congestion more likely and posing significant risks to traveler safety [2]. Therefore, understanding highway traffic flow patterns during holidays and achieving accurate, reliable traffic flow predictions are crucial for effective traffic management and control.

Recently, with the rapid development of artificial intelligent (AI), an increasing number of scholars have used deep learning-based methods to capture the spatial-temporal features for accurate short-term traffic prediction [3–5]. In 2014, Ma et al. [6] first applied the long short-term memory (LSTM) network to traffic flow prediction, achieving promising results. Since then, researchers have proposed numerous deep learning-based models for traffic flow prediction, filling the gap in its application within this field. For example, Zhang et al. [7] developed a deep learning-based model, ST-ResNet, which comprehensively models the temporal characteristics of crowd traffic, including closeness, periodicity, and trends,

thereby enabling accurate forecasting of both inbound flow and outbound flow dynamics in urban areas. Li et al. [8] regarded the traffic flow as a diffusion process on a directed graph and proposed a Diffusion Convolutional Recurrent Neural Network (DCRNN) to predict the traffic flow. Given the fact that traffic networks inherently possess a graph structure with rich topological information, numerous scholars have incorporated Graph Convolutional Networks (GCNs) into prediction models to effectively capture spatial dependencies. For example, Zhao et al. [9] applied GCNs to model the complex spatial dependencies in traffic flow, and utilized the Gated Recurrent Unit (GRU) to capture dynamic temporal dependencies, enabling accurate passenger flow prediction. Zhang et al. [10] proposed a GCN-based model, Conv-GCN, which utilized multi-graph convolution to model the spatial dependencies from recent, daily, and weekly perspectives. However, conventional GCNs model spatial dependencies using pre-defined graphs, which fail to account for their dynamic nature. To address this limitation, some researchers have extended GCNs to effectively capture dynamic spatial dependencies. For example, Wu et al. [11] developed a self-adaptive dependency matrix to accurately capture the hidden spatial dependencies in traffic data. Zhang et al. [12] constructed three types of graphs—the adjacency matrix graph, functional similarity graph, and OD correlation graph—to model the complex spatial dependencies of passenger flow. The attention mechanism [13] effectively captures the correlation between two nodes without considering their distance, addressing the long-range dependency problem in time series data. As a result, it has been widely applied to time series data processing in recent years [14–16]. For instance, Yan et al. [17] proposed a Traffic Transformer model consisting of a global encoder and a global–local encoder, leveraging the attention mechanism to effectively learn the dynamic multi-layer spatiotemporal features in traffic flow data. Xu et al. [18] introduced a novel Spatiotemporal Transformer Network (STTN), which utilizes multi-head attention to capture dynamic directional spatial dependencies and long-term temporal dependencies, enabling long-term traffic flow prediction. Zhang et al. [19] employed the attention mechanism to model temporal interactions among real-time, daily, and weekly demand, effectively modeling the complex temporal characteristics of traffic demand. In addition to the aforementioned short-term traffic flow prediction studies, several scholars have also explored long-term traffic flow prediction tasks [20,21]. The primary objective is to analyze the evolution trends of traffic over extended periods (exceeding one day), thereby supporting the development of more effective comprehensive management strategies. Wang et al. [22] proposed a hard attention mechanism based on learning similar patterns to enhance neuronal memory and reduce the accumulation of error propagation, accurately learning local feature and long-term dependence. Li et al. [23] developed a nonparametric performance-oriented prediction interval (PI) construction approach based on an enhanced sequential convolutional long short-term memory units (ConvLSTM) model, which can learn the long-term temporal correlations involved in the multivariate explanatory samples.

Although the above studies have effectively improved traffic flow prediction accuracy, most of them focus on urban road traffic or urban rail transit passenger flow, with relatively few studies addressing highway traffic flow prediction. Meanwhile, most research focuses on traffic flow prediction under normal scenarios, with limited studies on traffic flow prediction in abnormal scenarios, such as during holidays [24,25]. Holiday highway traffic flow exhibits significant uncertainty and suddenness, which may easily lead to traffic congestion and pose risks to the safety of travelers' lives and property. Additionally, in holidays, which span multiple days, traffic managers require accurate long-term traffic flow forecasts (beyond a single day) to formulate more effective management strategies. Therefore, this study investigates the long-term traffic flow forecasting during holidays.

To address the challenges outlined above, this study proposes a deep learning-based model for multi-step traffic flow prediction at highway toll stations during holidays. Given the interdependence between inbound and outbound flows at these stations, modeling them separately may fail to fully capture their mutual influence. Multi-task learning (MTL) has garnered significant attention in traffic flow prediction recently [26,27], owing to its capacity to enhance the generalization ability of models for various tasks by simultaneously learning multiple related tasks within a model and sharing feature representation knowledge. For example, Yi et al. [28] proposed a Continuous Multi-task Spatio-Temporal learning framework (CMuST) to not only reinforce individual correlated learning task in the collective perspective, but also help understand the cooperative mechanism of dynamic spatiotemporal systems. Yang et al. [29] proposed a multi-task-learning-based model called MultiMode-former (M2-former) to predict the network-wide short-term inflow of the multi-traffic modes system (metro, bus, and taxi). Zou et al. [30] developed a novel multi-task spatiotemporal network for highway traffic flow prediction (MT-STNet). By dividing the highway traffic flow prediction into three tasks and sharing the underlying traffic patterns and knowledge learned, their approach enhances the prediction performance of each subtask.

Inspired by the strong knowledge-sharing ability of multi-task learning, this study adopts a multi-task framework that jointly models inbound and outbound traffic flows by sharing the underlying traffic patterns, aiming to improve forecasting accuracy for both. To fully capture the spatial-temporal dependencies of inbound and outbound flow, this study develops a spatial-temporal cross-attention network (ST-Cross-Attn), which consists of bidirectional convolutional LSTM (Bi-ConvLSTM) and cross-attention (Cross-Attn) mechanism. Specifically, several identical Bi-ConvLSTM is used to capture the complex spatial-temporal dependencies of inbound and outbound flow, respectively. The Cross-Attn sequentially fuses the spatial-temporal dependencies of inbound and outbound flow, enabling spatial-temporal information sharing between both two and obtaining enhanced spatiotemporal features. Finally, based on the ST-Cross-Attn, this study develops a sequence-to-sequence (Seq2Seq) multi-step prediction framework. The encoder extracts enhanced feature encodings for both inbound and outbound flows at the current time step, while the decoder iteratively predicts traffic flows for future time steps. The proposed model is applied to a real-world traffic flow dataset from toll stations during the Labor Day holiday on highways in Shandong, China. The experimental results demonstrate that the proposed model achieves favorable performance in forecasting both inbound and outbound traffic flows at highway toll stations during holidays. Extensive ablation studies were conducted to demonstrate the advantages of jointly considering inbound and outbound flows and to analyze the interaction mechanisms among these flows. The contributions of this study are summarized as follows.

- (1) This study employs a multi-task learning method to achieve the joint prediction of inbound and outbound highway traffic flow. Compared to separately forecasting inbound and outbound flow at highway toll stations, the multi-task learning framework enables the sharing of underlying spatiotemporal patterns between inbound and outbound flows, which helps the model better capture their latent feature representations and improves overall prediction performance.
- (2) This study proposes a novel spatial-temporal cross-attention network (ST-Cross-Attn) based on the bidirectional ConvLSTM (Bi-ConvLSTM) and the cross-attention mechanism (Cross-Attn). The proposed model not only captures the spatial-temporal dependencies of inbound and outbound flow, but also facilitates the interaction between their underlying spatial-temporal features.

- (3) This study employs a sequence-to-sequence (Seq2Seq) multi-step prediction framework, in which the encoder progressively captures the hidden spatiotemporal features of inbound and outbound flow at the current time step, while the decoder iteratively generates multi-step prediction results. Extensive experiments on a real-world highway traffic flow dataset demonstrate that the proposed model outperforms state-of-the-art baselines in multi-step prediction of inbound and outbound flows at toll stations during holidays. The results of the ablation studies further demonstrate the effectiveness of the multi-task learning framework in predicting highway inbound and outbound traffic flows.

2. Preliminaries

This study aims to jointly predict the inbound and outbound flow of highway toll stations for multiple future time step during holidays. Leveraging historical traffic flow data from toll stations during holidays, this study extracts the inbound and outbound flow time series data at a 60 min time granularity to achieve joint multi-step prediction of highway inbound and outbound flow. Several fundamental concepts are first defined to formulate the highway traffic flow prediction problem.

Definition 1. (*Inbound/Outbound Flow Time Series Matrix*): The raw highway traffic flow data from toll stations are collected at a 60 min time granularity, including the toll station ID, time interval, inbound flow, and outbound flow data. Taking the inbound flow from toll stations as an example, let $p_{in}(n, t)$ denote the inbound flow observation at toll station n during the t -th time interval. Consequently, the inbound flow time series matrix for the toll station can be represented as follows,

$$p_{in} = \begin{pmatrix} p_{in}(1, 1) & p_{in}(1, 2) & \cdots & p_{in}(1, t) \\ p_{in}(2, 1) & p_{in}(2, 2) & \cdots & p_{in}(2, t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{in}(n, 1) & p_{in}(n, 2) & \cdots & p_{in}(n, t) \end{pmatrix} \quad (1)$$

where $p_{in} \in \mathbb{R}^{N \times T}$ denotes inbound flow time series matrix at highway toll stations throughout the period, N denotes the number of highway toll stations, and T represents the number of time steps throughout the period. Similarity, the outbound flow time series matrix of highway toll stations can be formulated as follows,

$$p_{out} = \begin{pmatrix} p_{out}(1, 1) & p_{out}(1, 2) & \cdots & p_{out}(1, t) \\ p_{out}(2, 1) & p_{out}(2, 2) & \cdots & p_{out}(2, t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{out}(n, 1) & p_{out}(n, 2) & \cdots & p_{out}(n, t) \end{pmatrix} \quad (2)$$

Definition 2. (*Augmented Inbound/Outbound Flow Time Series Matrix*): To effectively learn the holiday characteristics of inbound/outbound flow of toll stations, this study utilizes both historical holiday and pre-holiday traffic flow to predict the traffic flow for the next k time steps during the upcoming holiday. Specifically, the historical traffic flow data refers to the traffic flow observations from the same time period during the previous year's holiday. Let the current time be t , and define T as the total number of time steps for the pre-holiday traffic data. The augmented time series matrix for holiday traffic flow can be formulated as shown in Equation (3),

$$P_{in}(t) = \begin{pmatrix} p_{in}^{last}(1, t+1) & p_{in}^{last}(1, t+2) & \cdots & p_{in}^{last}(1, t+k) \\ p_{in}^{last}(2, t+1) & p_{in}^{last}(2, t+2) & \cdots & p_{in}^{last}(2, t+k) \\ \vdots & \vdots & \ddots & \vdots \\ p_{in}^{last}(n, t+1) & p_{in}^{last}(n, t+2) & \cdots & p_{in}^{last}(n, t+k) \end{pmatrix}, \begin{pmatrix} p_{in}(1, t-1) & p_{in}(1, t-2) & \cdots & p_{in}(1, t-T) \\ p_{in}(2, t-1) & p_{in}(2, t-2) & \cdots & p_{in}(2, t-T) \\ \vdots & \vdots & \ddots & \vdots \\ p_{in}(n, t-1) & p_{in}(n, t-2) & \cdots & p_{in}(n, t-T) \end{pmatrix} \quad (3)$$

where p_{in}^{last} represents the historical holiday traffic flow data, k denotes the number of future time steps to be predicted, while T denotes the total number of time steps for the pre-holiday traffic flow.

Holiday Traffic Flow Prediction Task: During officially designated holiday periods, given the augmented inbound and outbound flow time series matrices $P_{in}(t)$ and $P_{out}(t)$ from highway toll stations at the t -th time step, the holiday traffic flow prediction task aims to jointly predict the inbound and outbound flow for the next k time steps of the holiday, which can be formulated as follows,

$$[Y_{in}(t+1) \cdots Y_{in}(t+k), Y_{out}(t+1) \cdots Y_{out}(t+k)] = f(P_{in}(t), P_{out}(t)) \quad (4)$$

where f represents a learned highway traffic flow prediction model, $Y_{in}(\cdot)$ and $Y_{out}(\cdot)$ denotes the predicted future inbound and outbound flow, respectively.

3. Methodology

To fully capture the complex spatial-temporal dependencies of inbound and outbound flow from highway toll stations during holidays, this study proposes a novel spatial-temporal cross-attention network (ST-Cross-Attn), mainly consisting of three components: the inbound flow branch, the outbound flow branch, and the cross-attention interaction branch. The overall structure of ST-Cross-Attn is illustrated in Figure 1, whose core idea is to enhance the feature representations of inbound and outbound flow by facilitating the interaction of their spatial-temporal information. Specifically, the inbound flow branch processed the inbound flow data using a bidirectional convolutional LSTM (Bi-ConvLSTM) to fully capture the spatial-temporal dependencies from both forward and backward directions. Similarly, the outbound flow branch handles the outbound flow data, leveraging Bi-ConvLSTM to effectively extract the spatial-temporal dependencies. Finally, the outputs of these two branches are fed into the cross-attention interaction branch, which learns the interaction between inbound and outbound flow, enabling accurate correlation modeling and feature representations. Building upon the proposed ST-Cross-Attn, this study further develops a sequence-to-sequence (Seq2Seq) prediction framework for highway toll station traffic flow, aiming to jointly predict the inbound and outbound flow for the next k time steps. The following sections provide a detailed introduction to the components of the ST-Cross-Attn, and the multi-task learning-based prediction framework based on the ST-Cross-Attn.

3.1. The Inbound/Outbound Flow Branch

To accurately predict the traffic flow of highway toll stations for the next k time steps during the holidays, this study constructed augmented inbound and outbound flow time series matrices, $P_{in}(t)$ and $P_{out}(t)$, as inputs to the respective inbound and outbound flow branches. These matrices incorporate historical holiday traffic flow data, enabling the model to capture holiday-specific traffic patterns effectively.

Existing studies [31,32] have demonstrated the effectiveness of convolutional LSTM (ConvLSTM) [33] in modeling the spatial–temporal dependencies of time series data. Specifically, the ConvLSTM can be formulated as follows,

$$\begin{aligned}
 i_t &= \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \\
 f_t &= \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \\
 C_t &= f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \\
 o_t &= \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_o) \\
 H_t &= o_t \circ \tanh(C_t)
 \end{aligned} \tag{5}$$

where $X_1, X_2, \dots, X_t$ denote the model inputs, $C_1, C_2, \dots, C_t$ denote the cell outputs $H_1, H_2, \dots, H_t$ denotes the hidden states, and i_t, f_t, o_t represent the input gate, forget gate, and output gate, respectively. W_{**} represents the weight of convolutional operator, b_* represents the bias, σ denotes the sigmoid function, $\tanh(\cdot)$ denotes the hyperbolic tangent function, $*$ denotes the convolutional operator, and $\circ$ denotes Hadamard product.

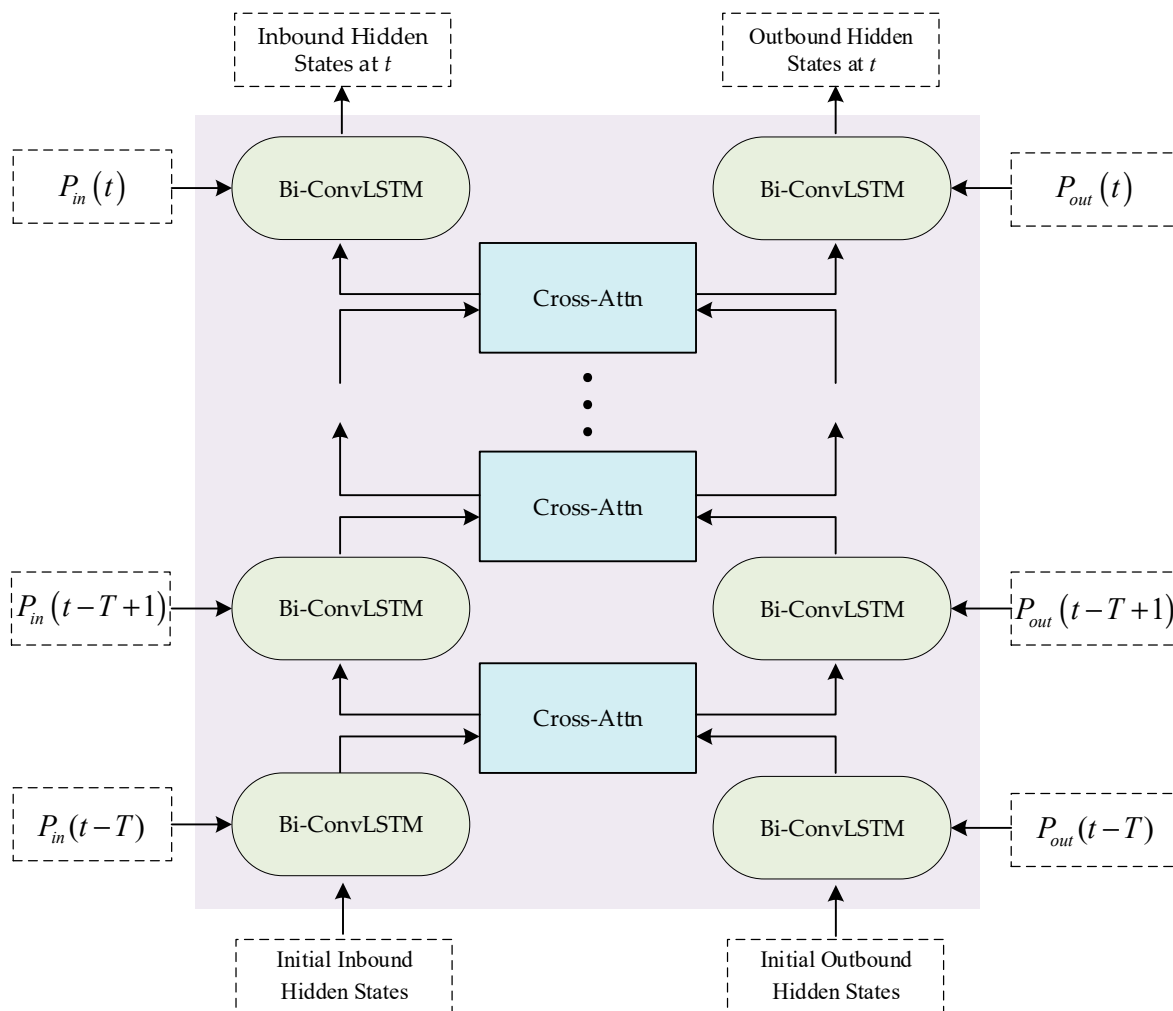


Figure 1. The architecture of ST-Cross-Attn.

However, unconventional time series data exhibit randomness and sudden fluctuations, making it challenging for ConvLSTM to effectively model complex spatial–temporal dependencies. This limitation leads to suboptimal performance in predicting unconventional time series data. A deeper analysis of such data reveals that, in unconventional scenarios like holidays, the forward time series maintains a certain level of continuity,

while the backward time series retains a degree of traceability. To fully capture the spatial-temporal characteristics of time series data in uncertain scenarios, this study introduces a bidirectional ConvLSTM (Bi-ConvLSTM), which can be shown as Figure 2.

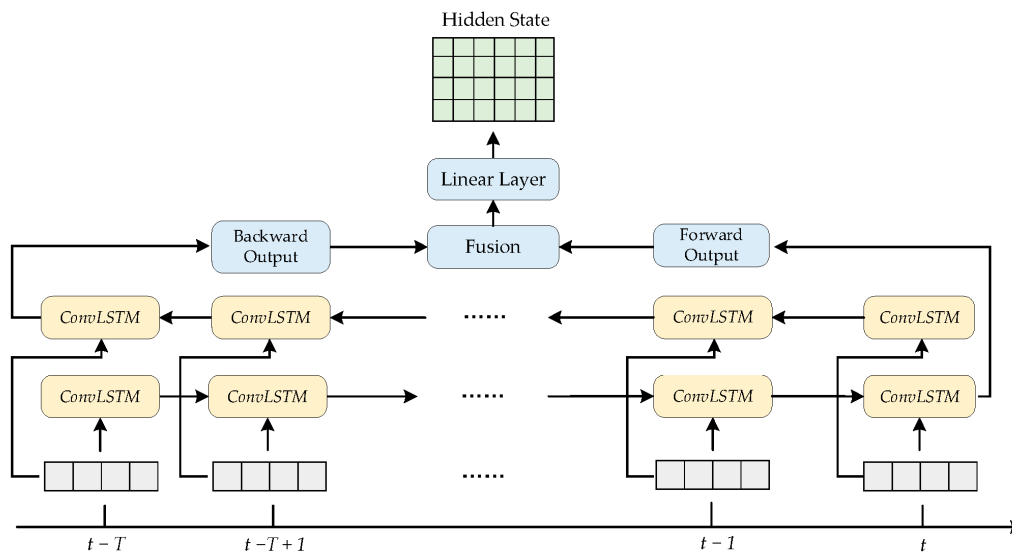


Figure 2. The architecture of Bi-ConvLSTM.

Specifically, the pre-processed inbound flow data $P_{in}(t) \in \mathbb{R}^{N \times (k+T)}$ are first dealt with by Bi-ConvLSTM in both forward and backward directions to comprehensively capture the spatial-temporal dependencies of inbound flow. Then the spatial-temporal dependencies captured by the forward and backward ConvLSTM are concatenated and further fused through a fully connected layer to obtain the final complex inbound hidden states $H_{in}(t) \in \mathbb{R}^{N \times (k+T)}$. Similar to the inbound flow, the outbound flow of highway toll stations also exhibits complex and dynamic spatial-temporal dependencies. Thus, the outbound flow branch also employs the Bi-ConvLSTMs to model the spatial-temporal dependencies from both forward and backward directions, ultimately obtaining the complex outbound hidden states $H_{out}(t) \in \mathbb{R}^{N \times (k+T)}$. Additionally, layer normalization is applied to accelerate training and improve network stability. Residual connections are used to mitigate issues such as gradient vanishing or exploding during model training.

3.2. The Cross-Attention Branch

Existing studies have extensively explored the relationship between metro inbound and outbound flow, revealing a strong correlation between them [34]. Similarly, the inbound and outbound flow at highway toll stations also exhibit a certain degree of causality and interdependence. Specifically, the inbound flow of a toll station at the current time period may influence the outbound flow at other toll stations in the future time period. This study defines this relationship as the causality between inbound and outbound flow. During holidays, this relationship is primarily driven by travel and leisure activities. For instance, travelers departing from a toll station to visit tourist attractions may lead to a surge in outbound flow at another toll station. Additionally, a strong correlation exists between the inbound and outbound flow at the same toll station. For example, in urban commuting scenarios, toll station traffic often exhibits tidal patterns. During the morning peak hours, the inbound flow at toll stations near residential areas are significantly higher than the outbound flow, whereas the reverse is observed during evening peak hours. Overall, there is a wealth of mutual information between toll station inbound and outbound flow. Effectively leveraging this interdependence for joint prediction can significantly enhance accuracy.

To this end, this study adopted the multi-task learning to realize the joint forecasting for inbound and outbound flow. Specifically, given the hidden states of inbound and outbound flow, this study further explores their interrelationship. Thus, a novel cross-attention branch is designed to capture the correlation between inbound and outbound flow using an attention mechanism. This mechanism facilitates the propagation of mutual information between the hidden states of inbound and outbound flow, effectively modeling the intrinsic interaction process between toll station inbound and outbound flow. The attention mechanism [13] is initially proposed for natural language processing (NLP) tasks, enabling the accurate computation of correlations between two nodes regardless of their distance. Due to its strong capability in feature extraction, this study improves the traditional attention mechanism and proposes a cross-attention module to jointly model the spatial-temporal features of inbound and outbound flow. The framework of the Cross-Attn module is illustrated in the Figure 3, which consists of two interactive attention submodules, the right submodule propagates the spatial-temporal information of the inbound flow to the outbound flow branch, while the left submodule propagates the spatial-temporal information of the outbound flow to the inbound flow branch. Taking the interaction process between the hidden states of the inbound flow branch $H_{in}(t)$ and the outbound flow branch $H_{out}(t)$ at time step t as an example, the computational process of the attention of the cross-attention branch is detailed as follows. First, the inbound hidden states and the outbound hidden states are processed by the embedding layer to obtain the Query, Key, and Value representations, respectively.

$$\begin{aligned} Q_{in}(t) &= \text{Conv}_{1 \times 1}(H_{in}(t), W_{in}^q), & Q_{out}(t) &= \text{Conv}_{1 \times 1}(H_{out}(t), W_{out}^q), \\ K_{in}(t) &= \text{Conv}_{1 \times 1}(H_{in}(t), W_{in}^k), & K_{out}(t) &= \text{Conv}_{1 \times 1}(H_{out}(t), W_{out}^k), \\ V_{in}(t) &= \text{Conv}_{1 \times 1}(H_{in}(t), W_{in}^v), & V_{out}(t) &= \text{Conv}_{1 \times 1}(H_{out}(t), W_{out}^v) \end{aligned} \quad (6)$$

where $\text{Conv}_{1 \times 1}$ represents the two-dimensional convolution operation with a 1×1 kernel, W_{*}^{*} represents the relevant weights. Query, key, and value representations all have the same dimensions as the hidden state of the inbound and outbound flow. Then, the information propagation coefficients for inbound flow $C_{in2out}(t)$ and outbound flow $C_{out2in}(t)$ are computed separately. These coefficients dynamically regulate the extent of information propagate between inbound and outbound flow.

$$\begin{aligned} C_{in2out}(t) &= \text{Softmax}(Q_{out}(t) \cdot K_{in}(t)^T), \\ C_{out2in}(t) &= \text{Softmax}(Q_{in}(t) \cdot K_{out}(t)^T) \end{aligned} \quad (7)$$

where $\text{Softmax}(\cdot)$ denotes the activation function and T denotes the matrix transpose operation. After obtaining the information propagation coefficients, the cross-attention of the inbound flow and outbound flow are calculated as follows.

$$\begin{aligned} \text{Attn}_{in2out}(t) &= C_{in2out}(t) \cdot V_{in}(t) \\ \text{Attn}_{out2in}(t) &= C_{out2in}(t) \cdot V_{out}(t) \end{aligned} \quad (8)$$

Given the cross-attention of inbound flow $\text{Attn}_{in}(t)$ and the cross-attention of outbound flow $\text{Attn}_{out}(t)$, the enhanced hidden states of inbound $\widehat{H}_{in}(t)$ and outbound flow $\widehat{H}_{out}(t)$ can be obtained by integrating the cross-attention with the hidden states through element-wise addition.

$$\hat{H}_{in}(t) = H_{in}(t) + Attn_{out \rightarrow in}(t)$$

$$\hat{H}_{out}(t) = H_{out}(t) + Attn_{in \rightarrow out}(t)$$

(9)

By propagating the mutual information of the inbound and outbound flow, the model allows each task to selectively focus on the most relevant features from the other task, which not only enhances their hidden states but also captures the causalities and correlations between them, thereby improving the accuracy of traffic flow predictions.

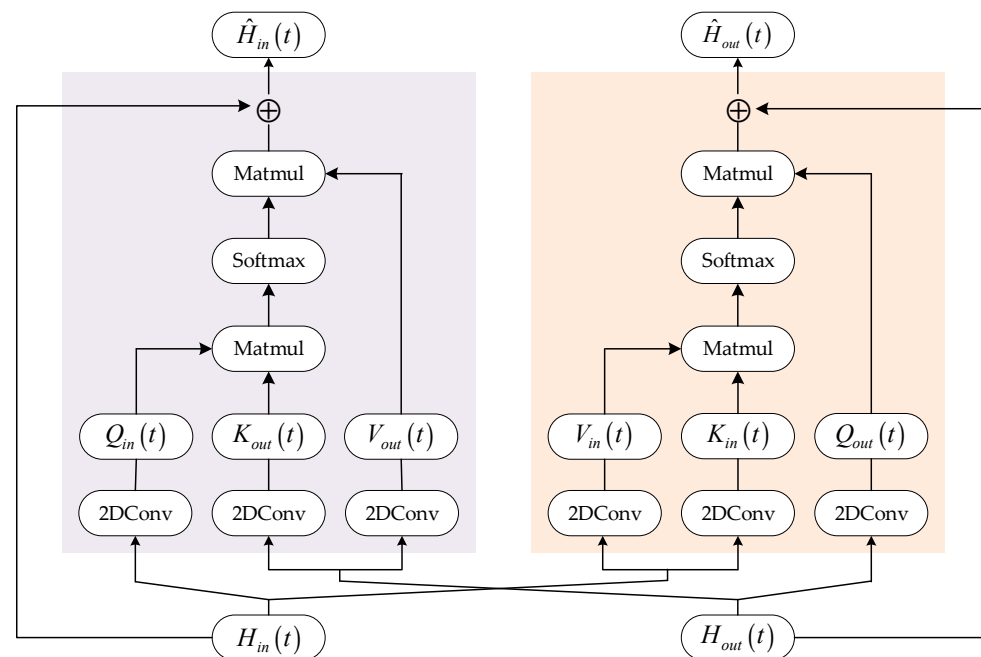


Figure 3. The structure of the cross-attention branch.

3.3. Sequence-to-Sequence (Seq2Seq) Prediction Framework

The proposed Bi-ConvLSTM effectively model the complex spatial-temporal dependencies of inbound and outbound flow of highway toll stations, while the Cross-Attn mechanism further capture the causalities and correlations between inbound and outbound flow, providing valuable mutual information for future predictions. Building on these two modules, this study introduces a unified multi-task learning-based prediction framework for toll station traffic flow, aiming to achieve simultaneous multi-step forecasting of inbound and outbound flow.

Specifically, the proposed framework follows a typical sequence-to-sequence (Seq2Seq) prediction architecture, where multi-step prediction is achieved through an encoder-decoder process. Both the encoder and decoder are designed as Bi-ConvLSTM, and the framework is illustrated in Figure 4. Taking the inbound flow as an example, the complete encoding and decoding process is introduced in detail, while the processing of the outbound flow follows the same procedure. Assuming the initial time is t , the inbound flow of all highway toll stations is denoted as $P_{in}(:, t)$, and the inbound hidden states from the previous time step is $H_{in}(:, t-1)$. These two inputs are processed by the Bi-ConvLSTM to encode the inbound hidden state at the current time step. Each hidden state at time t captures the historical spatial-temporal dependencies of the inbound flow. After obtaining the inbound hidden state $H_{in}(:, t)$ and the outbound hidden state $H_{out}(:, t)$, both of two are fed into the cross-attention branch to learn the causalities and correlations between them, so that generating the enhanced hidden states $\hat{H}_{in}(:, t)$ and $\hat{H}_{out}(:, t)$. Furthermore, the

obtained enhanced hidden states are used as the historical hidden states, and together with the inbound flow of the next time step, is fed into the Bi-ConvLSTM for future encoding. This process continues iteratively, obtaining the hidden state for the next time step and refining it through the cross-attention module to generate the enhanced hidden state. This iteration proceeds until the enhanced hidden state encoding for the current time step is obtained. It is important to note that at the initial time step, the inbound hidden state input to the Bi-ConvLSTM is initialized to zero. The above process constitutes the feature encoding for inbound traffic flow, which is identical to that of outbound traffic flow.

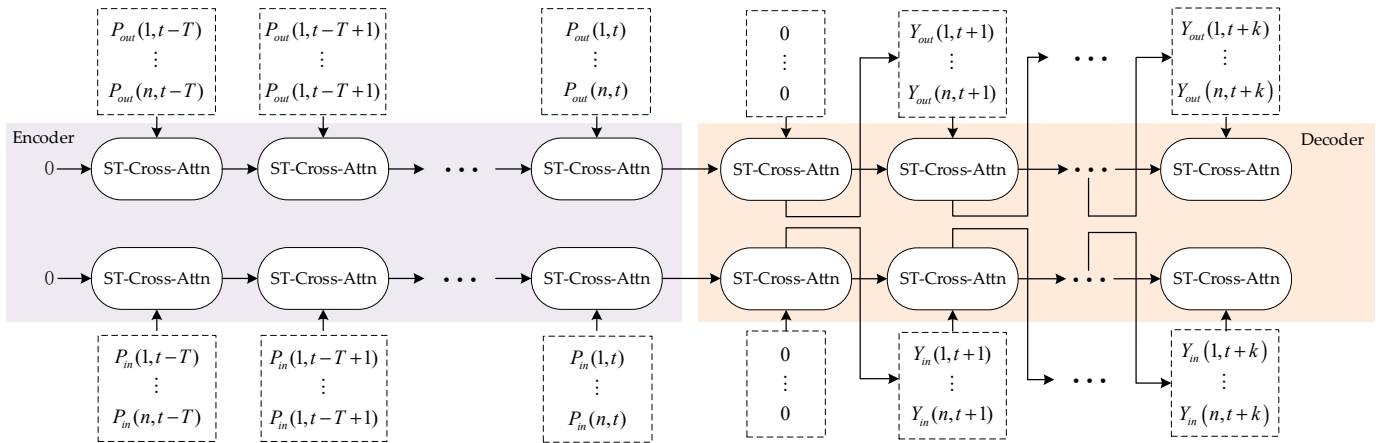


Figure 4. The structure of the sequence-to-sequence prediction framework.

Given the enhanced inbound hidden states at the current time step $\tilde{H}_{in}(:, t)$ from encoder, the ST-Cross-Attn within the decoder is used to generate the hidden state for the next time step $H_{in}(:, t+1)$. The fully connected layer then maps $H_{in}(:, t+1)$ to the corresponding feature representations $R_{in}(:, t+1)$. Finally, a residual connection is applied between $R_{in}(:, t+1)$ and the inbound flow at the same time from historical holiday data $P_{in}^{last}(:, t+1)$, yielding the predicted inbound flow $Y_{in}(t+1)$ for the next time step. Furthermore, the predicted inbound flow $Y_{in}(t+1)$ and the hidden state $H_{in}(:, t+1)$ at time $t+1$ are fed into the ST-Cross-Attn within the decoder to obtain the hidden state at time $t+2$. Subsequently, the predicted inbound flow at time $t+2$ is generated through a fully connected layer followed by a residual connection. By iterating the decoding process for T steps, the predicted inbound flow for the next T time steps $\{Y_{in}(t+1), Y_{in}(t+2), \dots, Y_{in}(t+T)\}$ can be obtained. Likewise, the same iterative process is applied to predict the outbound flow $\{Y_{out}(t+1), Y_{out}(t+2), \dots, Y_{out}(t+T)\}$ for the next T time steps.

4. Evaluation

4.1. Dataset

This study utilizes a real-world dataset containing inbound and outbound flow at highway toll stations in Shandong Province during the Labor Day holiday periods of 2021, 2023, and 2024. The 2022 data were excluded due to significant deviations caused by the COVID-19 pandemic. The dataset comprises records from 382 toll stations spanning ten days around Labor Day, specifically from 26 April to 5 May in each selected year. Each record includes the toll station name, station ID, date, hour, inbound flow, and outbound flow. To meet the input requirements of the model and mitigate inevitable data collection errors, the raw traffic data must be preprocessed before being used for prediction tasks. Specifically, during data cleaning, missing inbound or outbound flow values at certain toll stations are filled using mean interpolation. Records lacking corresponding station identifiers or timestamp information are discarded. Following data cleaning, inbound

and outbound flow information are extracted separately to construct time-series matrices, as defined in Equations (1)–(3). Following the definition of time series data in Section 2 (Preliminaries), the time granularity is set to 60 min. Subsequently, time series data for both toll station inbound and outbound flow are extracted accordingly. Notably, this study divides the traffic flow dataset into training, validation, and test sets in a 5:3:2 ratio.

To demonstrate that traffic patterns during the same holiday across different years exhibit similar trends, we use the Labor Day holiday as an example and compute the Pearson correlation coefficients of traffic flow between various years. The Pearson correlation coefficients are consistently above 0.7, which indicates a significant correlation in traffic flow during Labor Day holiday across different years. Based on this observation, we construct an augmented traffic flow matrix that integrates historical holiday data from multiple years. This approach enables the model to better capture the spatiotemporal characteristics of holiday traffic, thereby improving prediction accuracy. For instance, when forecasting traffic flow on 1 May 2023, the augmented matrix consists of two components: traffic flow data from 1 May 2021, and the most recent T time steps before the current time. Figure 5a,b respectively visualizes the inbound and outbound traffic flows during the Labor Day holidays across 2021 and 2023. It can be observed that the traffic patterns across the two years exhibit similar trends. Compared to regular working days, traffic volumes during the Labor Day holiday increased significantly and displayed more pronounced fluctuations.

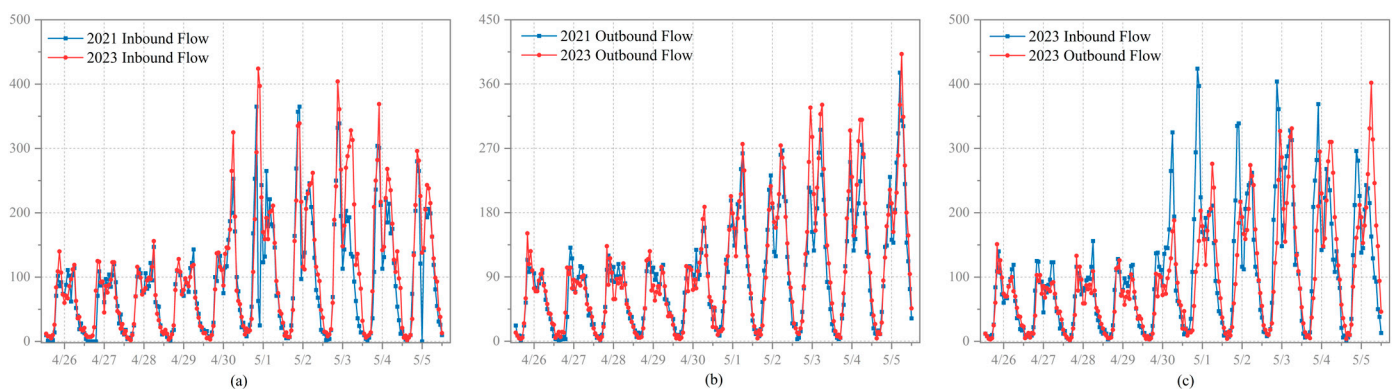


Figure 5. Comparison of inbound and outbound flow through highway toll stations during the Labor Day holiday. (a) Comparison of inbound flow during the 2021 and 2023 Labor Day holiday; (b) Comparison of outbound flow during 2021 and 2023 Labor Day holiday; (c) Comparison of inbound and outbound flow during 2023 Labor Day holiday.

Additionally, Figure 5c compares the inbound and outbound flow around the 2023 Labor Day holiday, revealing a clear contrast in their evolution patterns. While both flows increase significantly compared to regular working days, the inbound flow peaks on the first day of the holiday and then gradually declines, whereas the outbound flow steadily rises, reaching its peak on the final day. To capture these complementary evolution trends, this study employs a multi-task learning framework that jointly forecasts inbound and outbound flow. By sharing the underlying knowledge of traffic flow dynamics, the model can more effectively model the spatiotemporal features of inbound and outbound flow, thereby enhancing the overall prediction accuracy.

4.2. Model Settings and Evaluation Metrics

In this paper, all models are implemented with PyTorch 2.5.1 on a desktop computer with Intel® Core™ i9-10900X CPU, 64 GB memory, and an NVIDIA GeForce RTX3060 GPU.

Model settings: In the experiments, the input length of the model is set to 120 (i.e., 5 day). The ST-Cross-Attn model comprises two identical Bi-ConvLSTM layers and a

cross-attention branch. Each Bi-ConvLSTM layer has a hidden dimension of 512. In the cross-attention branch, the feature dimension is set to 512, and the number of attention heads is 8. All hyperparameters are fine-tuned to achieve optimal performance. Detailed settings are provided in Section 4.4. To improve training stability, dropout layers with a rate of 0.1 are applied to both the encoder and decoder. The model is trained using the Adam optimizer with a learning rate of 0.0001 and a batch size of 32.

Evaluation metrics: In the experiments, Mean Squared Error (MSE) is used as the loss function. Given the causal relationships and correlations between inbound and outbound flow, this study employs multi-task learning to joint predict both. The loss function is defined as the weighted sum of the MSE for inbound and outbound flow, with equal weights assigned to each, as shown in Equation (10).

$$Loss = MSE = \left[\frac{1}{n_{in}} \sum_{i=1}^{n_{in}} (y_{in}^i - \hat{y}_{in}^i)^2 \right] + \left[\frac{1}{n_{out}} \sum_{i=1}^{n_{out}} (y_{out}^i - \hat{y}_{out}^i)^2 \right], \quad (10)$$

Here, n_{in} and n_{out} represent the sample sizes of inbound and outbound flow, respectively. The y_{in}^i and y_{out}^i represent the actual value, and $\hat{y}_{in}^i$ and $\hat{y}_{out}^i$ represent the predicted value, respectively. To select an appropriate loss balancing strategy, we first train the model using equal loss weight coefficients for both tasks. The resulting loss values are then visualized to analyze the relative magnitudes of the two tasks, as shown in Figure 6.

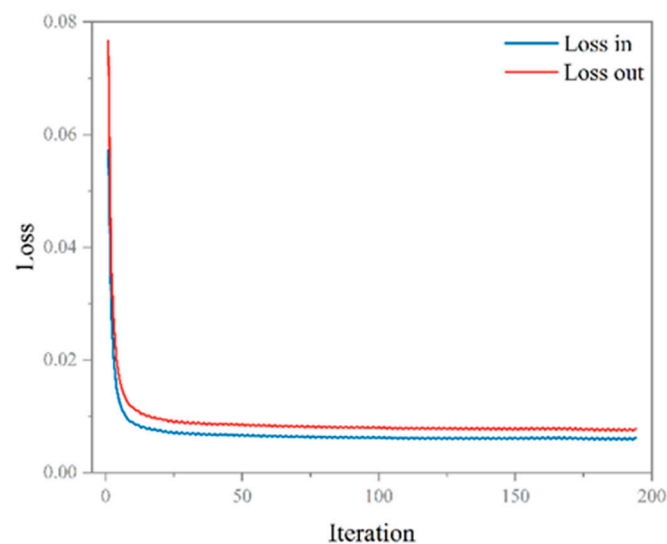


Figure 6. Curve graph of the model training loss function.

As shown in the visualization, although the loss of the inbound flow prediction task is generally lower than that of the outbound flow task, the two tasks have comparable loss scales and share similar structures, indicating that they are homogeneous. Therefore, we adopt a static weighted loss balancing strategy. Given the higher loss of the outbound flow prediction, it is designated as the dominant task. Specifically, we experiment with three sets of weight coefficients: [5, 5], [4, 6], and [3, 7]. The corresponding experimental results are summarized as Table 1.

Table 1. Prediction results under different loss weight coefficients (*1st Place*).

Weight Coefficients [Inbound, Outbound]	Inbound Flow			Outbound Flow		
	RMSE	MAE	WMAPE	RMSE	MAE	WMAPE
[0.5, 0.5]	139.3598	79.6379	0.3416	129.236	67.9112	0.3013
[0.4, 0.6]	115.167	65.2043	0.2796	95.164	52.5209	0.233
[0.3, 0.7]	119.34	70.7406	0.3033	102.788	62.0568	0.2753

As shown in the table, the model achieves the best prediction performance when the loss weights for the inbound and outbound flow prediction tasks are set to [0.4, 0.6]. Accordingly, this configuration is adopted throughout the experimental section. While the static weighting strategy is easy to implement, it may be suboptimal for heterogeneous tasks and often involves costly manual tuning. Future work will explore more adaptive loss balancing methods, such as GradNorm, to enhance training efficiency and task coordination.

This study uses Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Weighted Mean Absolute Percentage Error (WMAPE) as evaluation metrics for model performance. The formulas for these metrics are given in Equations (11)–(13). For simplicity, the actual values of traffic inbound and outbound flow are uniformly represented by y_i , the predicted values by $\hat{y}_i$, and the sample sizes for traffic inbound and outbound flow by n .

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (12)$$

$$WMAPE = \sum_{i=1}^n \left(\frac{y_i}{\sum_{j=1}^n y_j} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \quad (13)$$

4.3. Baselines

- (1) SVR (Support Vector Regression): As a classical machine learning model, Support Vector Regression (SVR) has been widely applied in the field of traffic forecasting [35], primarily utilizing Support Vector Machine techniques for regression analysis of time series data. In this experiment, the radial basis function kernel is used, with the error tolerance parameter set to $\epsilon = 0.005$ and the regularization parameter set to $C = 3$.
- (2) XGBoost (eXtreme Gradient Boosting) [36]: The Gradient Boosting Decision Tree algorithm is an efficient machine learning technique that combines multiple weak learners into a strong learner, thereby improving overall performance. In recent years, it has been widely applied in the field of traffic forecasting. In this study, the 'gbtree' booster is used, with a maximum tree depth of 6 and a learning rate of 0.1.
- (3) BPNN (Back Propagation Neural Network): As one of the most classic neural networks, the Back Propagation Neural Network (BPNN) has been shown to effectively capture the nonlinear characteristics of traffic flow [37]. The BPNN model in this study consists of two fully connected layers, with 1024 and 512 neurons in each layer, respectively. The optimizer used for the model is Adam, with a learning rate of 0.0005.
- (4) CNN (Convolutional Neural Network): Convolutional Neural Networks (CNNs) are known for their strong performance in capturing spatial correlations within data [38]. The convolutional kernels are effective in capturing the spatial dependencies between features in time series data. In this study, a CNN layer with a kernel size of 3×3 is first used to model the spatial features, followed by a fully connected layer for prediction, with ReLU as the activation function.
- (5) LSTM (Long Short-Term Memory): Long short-term memory (LSTM) networks effectively address the long-term dependency issue in conventional Recurrent Neural Networks (RNNs), enabling them to capture longer-term trends in time series data. As a result, LSTMs have certain advantages in traffic flow forecasting [39]. In this study, the model consists of one LSTM layer with four hidden layers and one fully connected layer, with ReLU as the activation function.
- (6) GRU (Gated Recurrent Unit): Gated Recurrent Units (GRUs) are a type of RNN that effectively solve the long-term dependency problem and the vanishing gradient issue

during backpropagation. GRUs have been widely applied in time series processing tasks [40]. Compared to LSTMs, GRUs feature a simpler structure, improving training speed while maintaining good performance. The model is composed of three hidden layers and one fully connected layer, with 128 neurons in each hidden layer.

- (7) ST-ResNet (Spatial Temporal Residual Network) [7]: This network uses residual convolutional units to model the spatiotemporal dependencies of traffic flow, employing three branches to learn the recent, periodic, and trend-related temporal features of traffic flow. In this study, a residual convolutional layer is used to learn spatiotemporal features, with a kernel size of 3×3 , and a fully connected layer is used for future traffic flow prediction.
- (8) Transformer: Initially proposed for Natural Language Processing (NLP) in 2017 [13], the Transformer model effectively captures correlations between any nodes without considering their distance. As a result, Transformer models have been widely applied in time series processing in recent years. In this study, a Transformer framework is constructed with two encoder layers and two decoder layers. The number of attention heads is set to 4, and the feature size d_{model} is set to 256. The output of the Transformer is then passed to a fully connected layer for prediction.
- (9) STMTL (Spatial–Temporal Multi-Task Learning) [25]: This model employs a Multi-Graph Channel Attention Network to capture both dynamic and static topological relationships in traffic flow, a Time-Encoding Gated Recurrent Unit (TE-GRU) to model the unique fluctuations during holidays, and an attention block to capture the interactive relationship between inflow and outflow. In this study, a dynamic similarity graph is constructed based on inflow and outflow data, while a static adjacency graph is not utilized.

Notably, these baseline models independently predict inbound and outbound flow using identical modules.

4.4. Study of Hyperparameters

To investigate the impact of different hyperparameters on model performance, this section focuses on three key parameters: input length, hidden dimension, and the number of attention heads. Specifically, the input length is set to {24, 48, 72, 96, 120}, the hidden dimension of the cross-attention module is set to {128, 256, 512}, and the number of attention heads is set to {2, 4, 8}. In the fine-tuning process, we follow the control variable principle, which means that only one hyperparameter will be tuned while the remaining hyperparameters remain constant until finding the best result.

Figure 7 illustrates the impact of the input length in prediction performance. It can be observed that when the input length is set to 120-time steps (i.e., five days), the proposed model achieves the best prediction performance across all prediction horizons. This suggests that incorporating a sufficiently long historical input enables the model to capture richer spatiotemporal dependencies, which are particularly important for forecasting complex holiday traffic dynamics. However, the results also indicate that a longer input does not always guarantee better performance. For instance, when the input length is 96-time steps (i.e., four day), the model's accuracy declines noticeably. This may be attributed to the introduction of redundant or noisy information, which could weaken the model's ability to focus on truly relevant patterns. Therefore, selecting an appropriate input length is essential for balancing information richness and model generalization. In addition, regardless of the input length, the prediction accuracy consistently decreases as the forecasting horizon extends. This is primarily due to the accumulation of errors inherent in multi-step forecasting.

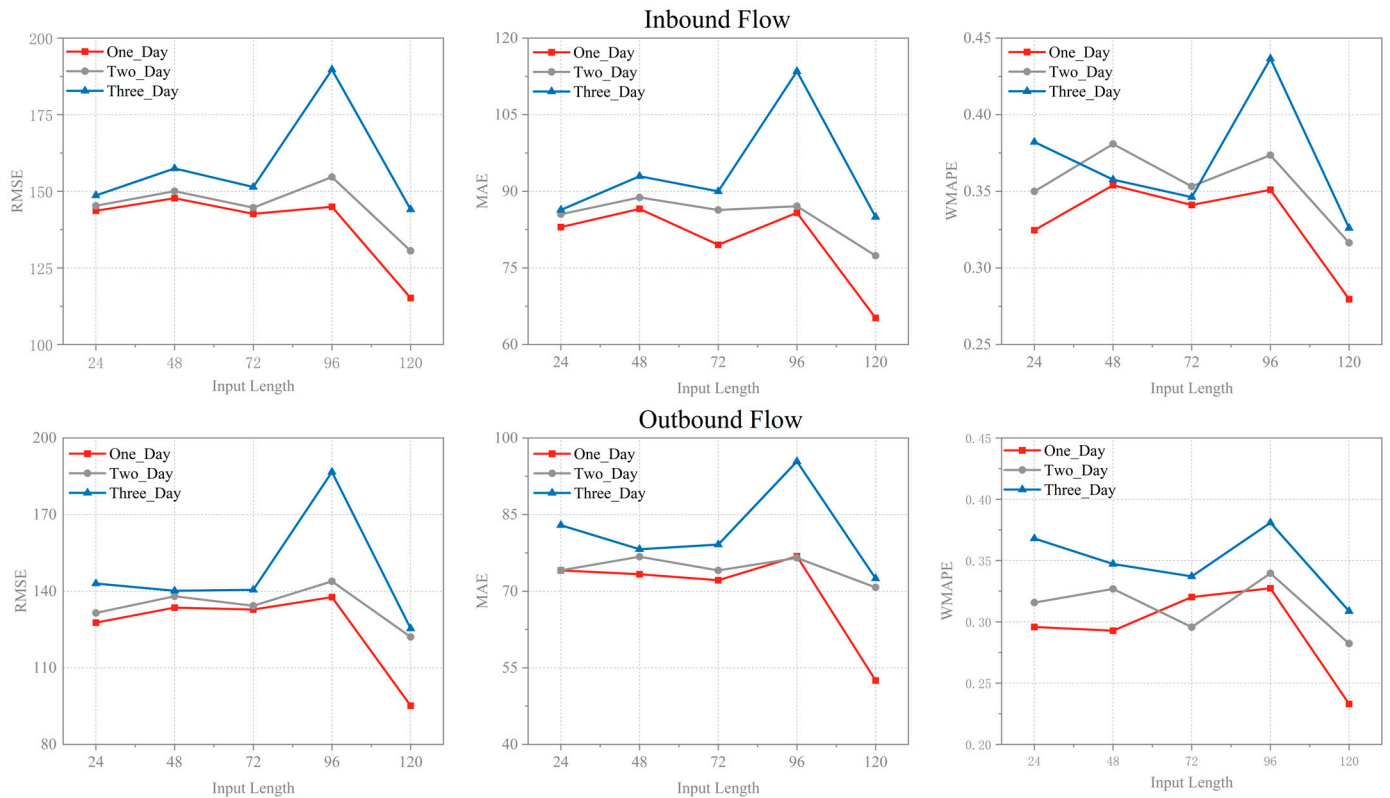


Figure 7. Impact of input length on model forecasting performance.

Figure 8 illustrates the impact of hidden dimensions and the number of attention heads on the model's performance. It is worth noting that these experiments were conducted under the setting of an input length of 120-time steps and a forecasting horizon of 24-time steps. As shown, the overall favorable prediction performance is achieved when the hidden dimension of the cross-attention mechanism is set to 512. Furthermore, when the hidden dimension is fixed at 512, the model reaches optimal performance for both inbound and outbound flow predictions when the number of attention heads is set to 8.

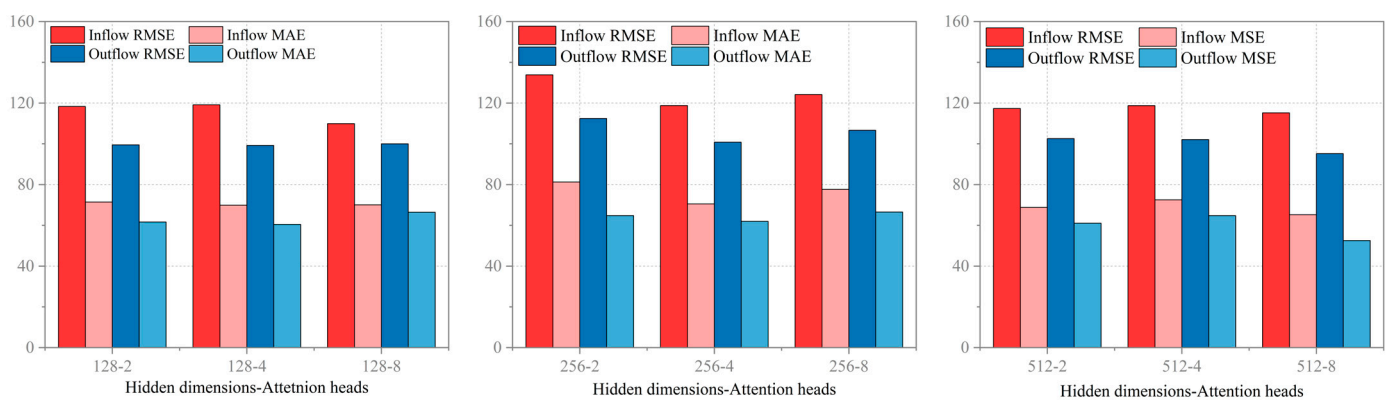


Figure 8. Impact of hidden dimensions and number of attention heads on model performance.

Overall, based on the experimental results, the optimal hyperparameter settings of the model are as follows: the input length is set to 120-time steps, the hidden dimension of the cross-attention mechanism is set to 512, and the number of attention heads is set to 8. All subsequent experiments were conducted using these hyperparameter configurations.

4.5. Network-Wide Performance Comparison

This section first analyzes the prediction results of all models in the real-world dataset of inbound and outbound flow from highway toll stations. Tables 2 and 3 present the average prediction results after five trials. Specifically, the machine learning-based SVR and XGBoost models demonstrate poorer predictive performance compared to other models across different prediction time steps. This is because when performing multi-step prediction, these models can only make independent predictions for each time step, failing to capture the long-term dependencies in the traffic flow time-series data. Thus, conventional machine learning-based models are not suitable for multi-step highway traffic flow prediction during holidays.

Table 2. Inbound flow prediction result evaluation (mean $\pm$ standard deviation, *1st Place*).

Model	1 Day (24 Time Steps)			2 Days (48 Time Steps)			3 Days (72 Time Steps)		
	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 $\pm$ 1.23	85.82 $\pm$ 0.83	33.06 $\pm$ 1.63%	191.70 $\pm$ 2.12	103.15 $\pm$ 1.24	39.35 $\pm$ 2.03%	187.66 $\pm$ 2.97	96.23 $\pm$ 1.01	34.56 $\pm$ 1.89%
CNN	158.84 $\pm$ 1.14	84.72 $\pm$ 0.85	30.77 $\pm$ 1.34%	166.27 $\pm$ 2.34	87.72 $\pm$ 0.99	31.51 $\pm$ 1.87%	184.91 $\pm$ 3.01	99.41 $\pm$ 1.25	35.73 $\pm$ 2.01%
LSTM	163.29 $\pm$ 1.11	84.88 $\pm$ 0.79	30.66 $\pm$ 1.45%	159.66 $\pm$ 2.01	83.48 $\pm$ 1.02	29.82 $\pm$ 2.10%	161.41 $\pm$ 3.32	88.69 $\pm$ 1.45	34.16 $\pm$ 2.54%
GRU	180.55 $\pm$ 1.05	94.12 $\pm$ 0.94	34.06 $\pm$ 1.22%	175.07 $\pm$ 2.44	90.44 $\pm$ 0.95	32.36 $\pm$ 1.94%	159.82 $\pm$ 3.43	87.30 $\pm$ 1.36	34.14 $\pm$ 2.28%
ST-ResNet	151.07 $\pm$ 0.99	79.28 $\pm$ 0.88	28.78 $\pm$ 1.56%	161.18 $\pm$ 1.90	85.18 $\pm$ 1.16	30.58 $\pm$ 2.05%	180.75 $\pm$ 3.22	99.57 $\pm$ 1.39	35.76 $\pm$ 2.06%
Transformer	145.81 $\pm$ 1.09	87.24 $\pm$ 0.76	31.68 $\pm$ 1.44%	148.63 $\pm$ 1.98	86.34 $\pm$ 1.21	30.98 $\pm$ 1.99%	153.32 $\pm$ 3.37	86.37 $\pm$ 1.28	33.86 $\pm$ 2.41%
STMTL	151.70 $\pm$ 1.17	78.66 $\pm$ 0.97	28.94 $\pm$ 1.33%	163.92 $\pm$ 1.90	84.10 $\pm$ 1.09	30.20 $\pm$ 1.89%	184.12 $\pm$ 3.01	93.72 $\pm$ 1.18	33.67 $\pm$ 2.21%
Our Model	115.17 $\pm$ 1.01	65.20 $\pm$ 0.81	27.96 $\pm$ 1.39%	130.56 $\pm$ 2.01	77.39 $\pm$ 1.13	30.11 $\pm$ 1.91%	144.08 $\pm$ 3.21	84.95 $\pm$ 1.30	32.47 $\pm$ 2.25%

Table 3. Outbound flow prediction result evaluation (mean $\pm$ standard deviation, *1st Place*).

Model	1 Day (24 Time Steps)			2 Days (48 Time Steps)			3 Days (72 Time Steps)		
	RMSE	MAE	WMAPE	RMSE	RMSE	MAE	WMAPE	MAE	RMSE
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 ± 1.31	75.65 ± 0.99	28.41 ± 1.34%	158.47 ± 2.17	86.28 ± 1.31	33.53 ± 2.15%	152.34 ± 3.34	81.43 ± 1.66	29.87 ± 1.98%
CNN	126.68 ± 1.24	69.58 ± 0.93	25.69 ± 1.28%	145.30 ± 2.09	81.07 ± 1.22	29.74 ± 2.07%	161.56 ± 3.58	94.59 ± 1.82	34.69 ± 2.33%
LSTM	133.85 ± 1.29	69.04 ± 0.88	25.49 ± 1.30%	132.42 ± 2.01	75.41 ± 1.36	25.45 ± 2.33%	139.51 ± 3.49	72.53 ± 1.69	29.61 ± 2.03%
GRU	140.69 ± 1.14	73.59 ± 0.92	27.18 ± 1.27%	136.94 ± 1.99	73.15 ± 1.30	26.47 ± 2.18%	130.81 ± 3.62	73.20 ± 1.89	29.85 ± 2.01%
ST-ResNet	132.01 ± 1.25	71.03 ± 0.86	26.23 ± 1.37%	146.56 ± 2.10	82.07 ± 1.25	30.10 ± 2.31%	162.27 ± 3.44	92.54 ± 2.01	33.94 ± 2.16%
Transformer	117.49 ± 1.22	75.56 ± 0.83	27.89 ± 1.19%	127.39 ± 2.20	73.71 ± 1.33	25.94 ± 2.14%	132.44 ± 3.28	76.64 ± 1.92	28.38 ± 2.28%
STMTL	128.07 ± 1.12	70.09 ± 0.82	25.88 ± 1.03%	138.95 ± 2.00	75.16 ± 1.25	27.57 ± 1.98%	160.80 ± 3.15	84.59 ± 1.78	31.03 ± 2.01%
Our Model	95.16±1.19	52.52±0.90	23.33±1.25%	125.39±2.04	72.48±1.29	25.37±2.10%	122.07±3.31	70.74±1.74	28.22±2.19%

In terms of deep learning-based models, although the BPNN can capture the nonlinear relationships in holiday traffic flow, its inability to model the complex spatial–temporal dependencies results in a performance that only surpasses machine learning-based models. LSTM, GRU, and Transformer, deliver better prediction performance than BPNN by effectively capturing temporal dependencies. Transformer outperforms both LSTM and GRU by leveraging the self-attention mechanism to capture long-term dependencies. Although CNN can capture local spatial relationships and temporal dependencies through convolution operations, its predictive performance deteriorates as the number of time steps increases due to its limited ability to model long-term traffic dynamics. ST-ResNet, which models both temporal and spatial dependencies in a global context, achieves higher prediction accuracy than other models, but it fails to account for the interaction between traffic inbound and outbound flow. STMTL further improves prediction accuracy by incorporating the interaction between inflow and outflow as well as modeling the unique fluctuations associated with holidays. However, relying solely on the dynamic similarity graph without incorporating the physical adjacency graph limits the model’s ability to achieve satisfactory prediction performance.

Finally, our model employs Bi-ConvLSTM to simultaneously capture spatial dependencies and global temporal dependencies, while utilizing Cross-Attn to propagate mutual spatial–temporal features between inbound and outbound hidden states, enhancing fea-

ture representations. Moreover, our model leverages the augmented inbound/outbound flow time series matrices as the input to effectively learn the holiday characteristics of inbound/outbound flow of toll stations. Consequently, our model effectively models the spatial–temporal dependencies of inbound and outbound flow, enabling accurate joint prediction of highway toll station traffic during holidays.

4.6. Prediction Performance of Individual Toll Stations

During holidays, different toll station types exhibit varying holiday traffic flow characteristics. To verify the model’s prediction accuracy and robustness, we visualized and analyzed the forecasting results across different forecasting horizons and toll station types.

The first toll station, Mengyin Toll Station in Linyi City, shows a balanced traffic pattern during holidays, with traffic inbound and outbound flow being approximately equal. As shown in Figure 9, the inbound and outbound flow follows distinct morning and evening peak patterns on weekdays, where holiday traffic is significantly higher than weekday traffic. The model demonstrates a strong fit between predictions and actual values across different forecasting horizons, confirming its high predictive accuracy.

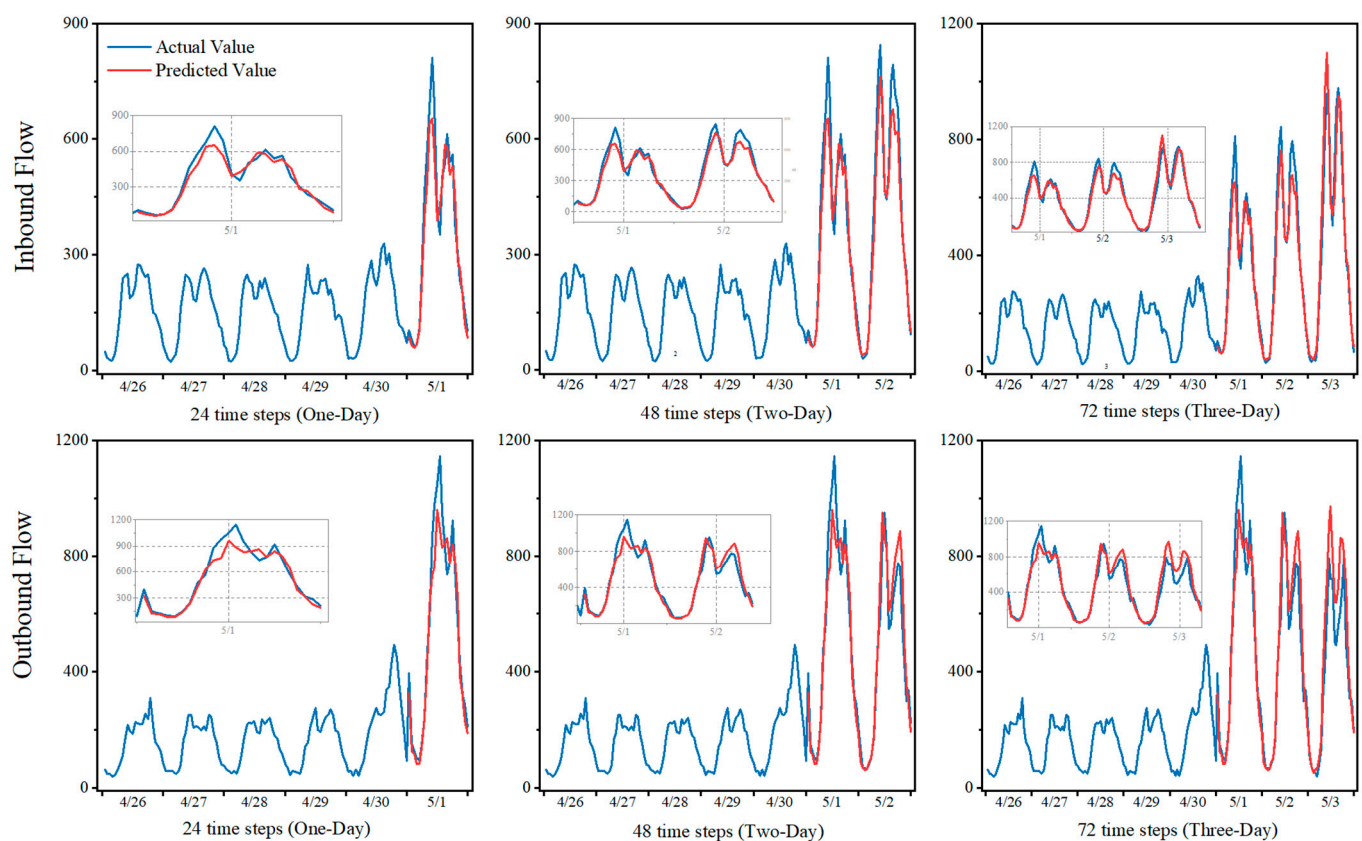


Figure 9. Prediction–actual value comparison curve at Mengyin Station.

The second toll station, Bucun Toll Station in Jinan City, is an inbound flow-dominant station, where inbound flow traffic exceeds outbound flow during holidays. As illustrated in Figure 10, this station exhibits clear weekday traffic patterns with pronounced morning and evening peaks. During holidays, the inbound flow is slightly higher than on weekdays, while the outbound flow remains largely unchanged, indicating a moderate increase in total traffic volume. The predicted values closely align with the actual data across all forecasting horizons, confirming the model’s effectiveness.

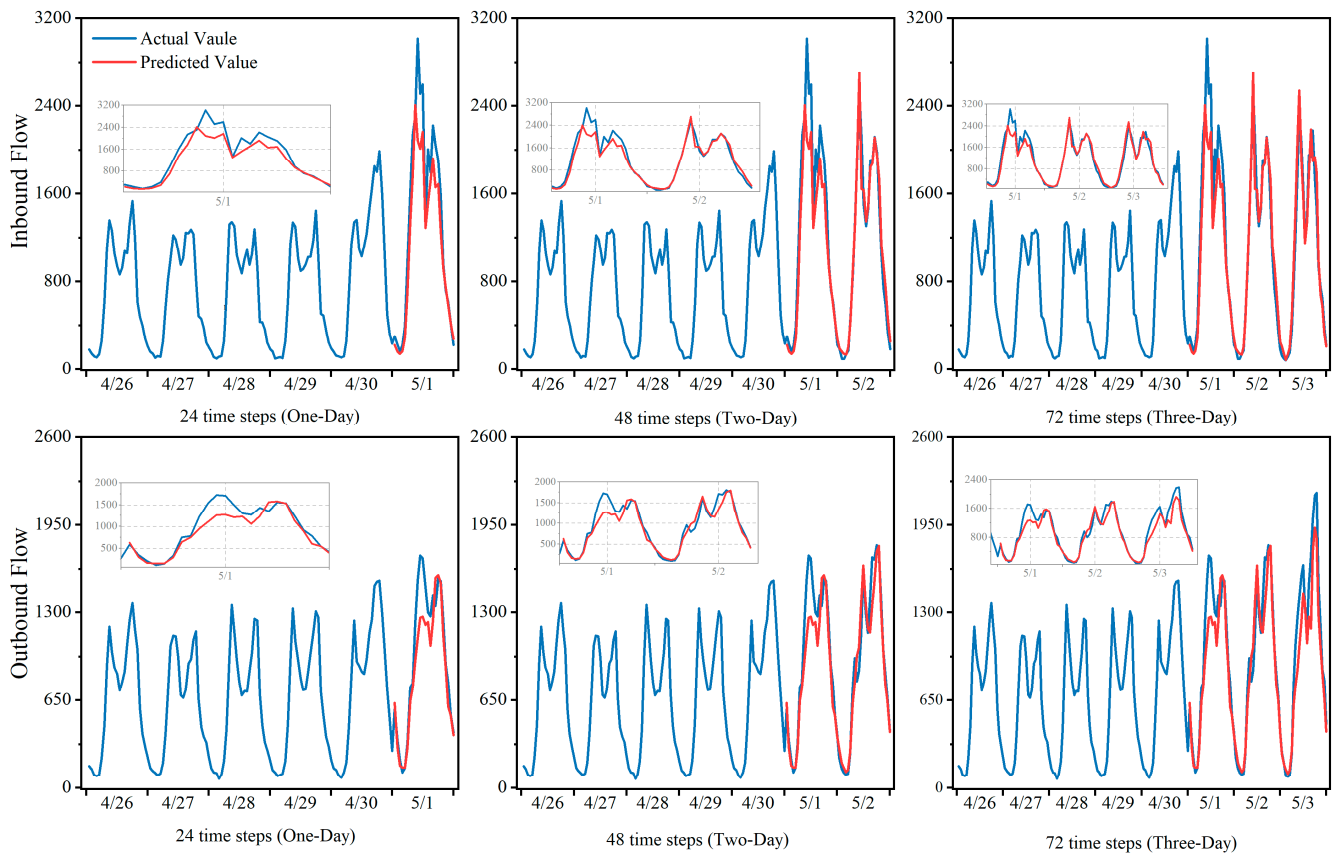


Figure 10. Prediction–actual value comparison curve at Bucun Station.

The third toll station, Pengji Toll Station in Tai'an City, is an outbound flow-dominant station, where traffic outbound flow exceeds traffic inbound flow during holidays. As shown in Figure 11, the station experiences lower traffic volumes on weekdays compared to the previous two toll stations. However, during holidays, both inbound and outbound flow increase significantly. Despite this surge, the proposed model effectively captures the variations in traffic flow patterns, demonstrating strong robustness and reliability.

4.7. Prediction Performance of Ablation Studies

To validate the effectiveness of each component within the proposed model, this section conducts extensive ablation experiments with the following model variants:

Variant 1: ST-Cross-Attn-No Bi: In this variant, the bidirectional convolutional long short-term memory (Bi-ConvLSTM) module in the spatiotemporal attention interaction model is replaced with a standard convolutional long short-term memory (ConvLSTM) module, while keeping all other configurations unchanged. This experiment aims to investigate the impact of bidirectional modeling on feature extraction.

Variant 2: ST-Cross-Attn-LSTM: This variant replaces the Bi-ConvLSTM module with a bidirectional long short-term memory (Bi-LSTM) module while keeping other configurations unchanged. The purpose of this experiment is to explore the capability of convolution operations in capturing spatiotemporal dependencies.

Variant 3: ST-Cross-Attn-No Attn: In this variant, the attention interaction module is removed to examine the role of attention interaction in enhancing feature representation. Notably, without this module, the hidden features of traffic inbound and outbound flow are no longer shared, effectively transforming the model into a simple combination of two single-task learning models. However, since the loss function still incorporates both

traffic inbound and outbound flow components, this model remains distinct from a purely single-task learning approach.

Variant 4: ST-Cross-Attn-No Holiday: To capture holiday traffic patterns, this study proposes an augmented time-series matrix that utilizes historical holiday traffic flow data to assist in predicting future holiday traffic flow. In this variant, the original traffic inbound and outbound flow time-series matrix is used instead, aiming to assess the contribution of historical holiday traffic flow data in improving prediction accuracy.

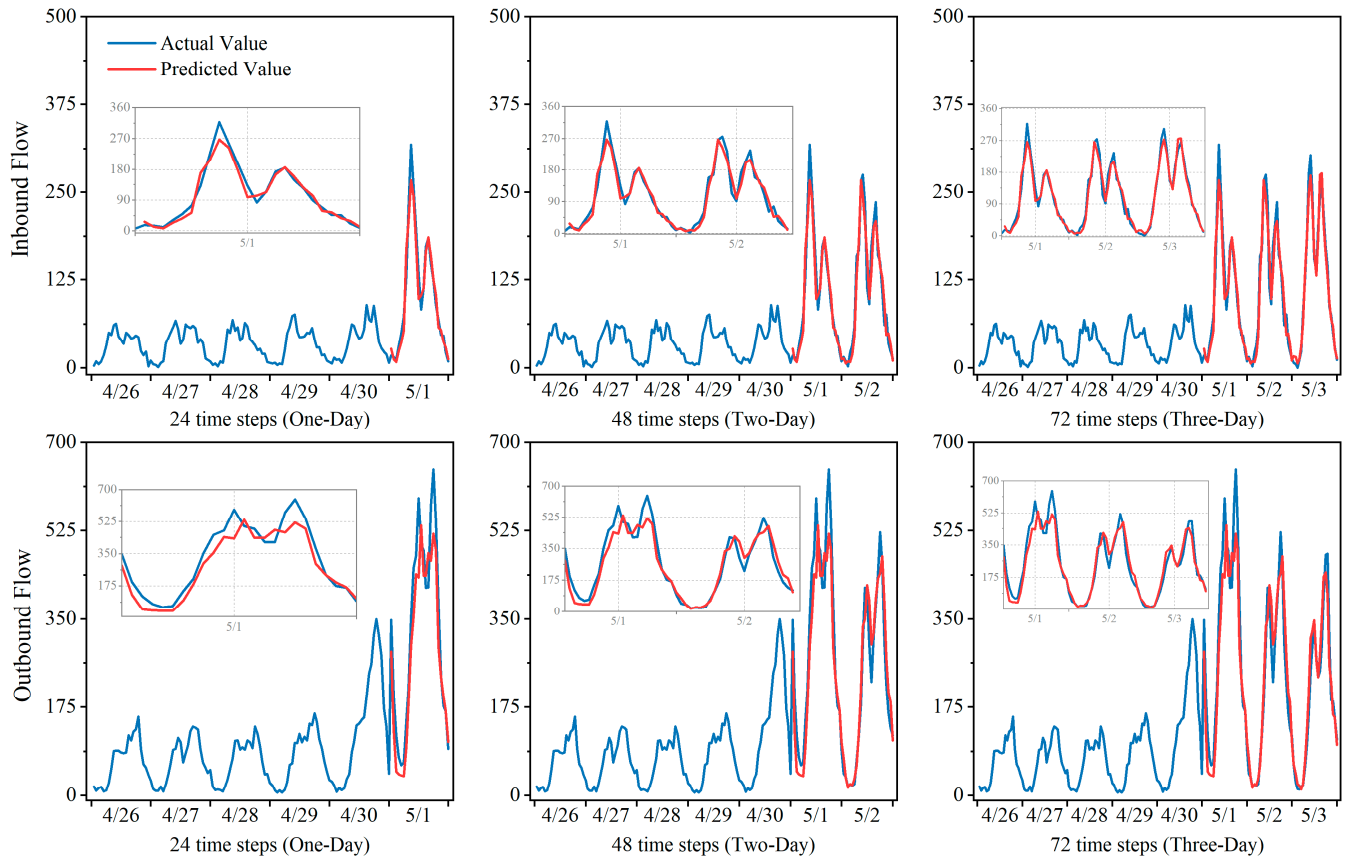


Figure 11. Prediction–actual value comparison curve at Pengji Station.

According to the results recorded in Tables 4 and 5, the following conclusions can be drawn from the ablation experiments:

Table 4. Inbound flow prediction result evaluation of variants (*1st Place*).

Variant	1 Day (24 Time Steps)			2 Days (48 Time Steps)			3 Days (72 Time Steps)		
	RMSE	MAE	WMAPE	RMSE	RMSE	MAE	WMAPE	MAE	RMSE
Variant 1 (No Bi)	137.55	68.56	29.25%	143.21	72.04	29.19%	155.94	87.61	33.42%
Variant 2 (LSTM)	152.92	83.85	35.87%	158.29	86.19	34.96%	160.47	85.26	32.53%
Variant 3 (No Attn)	139.84	74.35	31.71%	145.76	77.67	31.47%	169.76	83.35	31.80%
Variant 4 (No holidays)	165.83	76.84	32.10%	160.47	75.03	32.08%	153.39	74.12	30.01%
<i>ST-Cross-Attn</i>	136.89	66.91	28.55%	142.32	70.27	28.47%	148.34	72.08	27.47%

Table 5. Outbound flow prediction result evaluation of variants (*1st Place*).

Variant	1 Day (24 Time Steps)			2 Days (48 Time Steps)			3 Days (72 Time Steps)		
	RMSE	MAE	WMAPE	RMSE	RMSE	MAE	WMAPE	MAE	RMSE
Variant 1 (No Bi)	117.23	55.88	24.79%	124.22	61.72	26.87%	132.51	72.22	28.82%
Variant 2 (LSTM)	135.35	71.09	31.55%	139.64	73.63	31.36%	142.27	73.24	29.24%
Variant 3 (No Attn)	119.10	60.01	26.62%	125.71	64.83	27.61%	162.37	79.87	31.89%
Variant 4 (No holidays)	167.52	65.37	29.02%	154.16	72.71	30.83%	138.08	71.69	28.63%
<i>ST-Cross-Attn</i>	116.63	54.70	24.28%	122.86	59.54	25.37%	127.39	61.99	24.76%

Effectiveness of Bidirectional Modeling (Variant 1: ST-Cross-Attn-No Bi): This variant replaces Bi-ConvLSTM module with a unidirectional ConvLSTM, restricting the feature learning to one direction. The experimental results show a decline in prediction accuracy, indicating that unidirectional modeling inadequately captures temporal dependencies and causality in sequential data. Furthermore, as the prediction horizon increases, prediction error increases significantly, highlighting the superiority of bidirectional modeling for multi-step forecasting. This suggests that bidirectional modeling effectively captures long-term spatiotemporal dependencies, mitigating the degradation of feature representation over extended time steps.

Impact of Convolutional Operations (Variant 2: ST-Cross-Attn-LSTM): In this variant, Bi-ConvLSTM is replaced with a bidirectional LSTM (Bi-LSTM). The results show a significant decline in prediction accuracy, indicating that, compared to ConvLSTM, LSTM fails to effectively model the spatial characteristics of traffic flow, thereby affecting the overall predictive performance. Highway toll station traffic flow exhibits complex spatial characteristics, such as the strong correlation between urban commuter toll stations, where traffic inbound and outbound flow often exhibit tidal patterns. The convolutional operators in ConvLSTM allows the model to capture these complex spatial dependencies, thereby improving the prediction accuracy.

Effect of Attention Interaction (Variant 3: ST-Cross-Attn-No Attn): This variant removes the attention interaction mechanism, eliminating the dependency between inbound and outbound flow. The results indicate that, compared to the ST-Cross-Attn, this variant exhibits lower prediction accuracy, with the decline becoming more pronounced as the prediction horizon increases. This occurs because attention mechanisms enable the model to capture long-term dependencies between spatial nodes, disregarding distance constraints. Without the attention interaction module, the model's ability to capture long-term correlations weakens, leading to significant performance degradation in long-term prediction. Additionally, the attention interaction module facilitates the sharing of spatial-temporal information between inbound and outbound flow, enhancing feature representation. Removing this module disrupts the information propagation, preventing effective spatial-temporal feature interaction and resulting in lower prediction accuracy. Therefore, the Cross-Attn not only captures dependencies between long-term time-series nodes but also enhances feature learning by enabling information propagation between inbound and outbound flow, ultimately improving prediction performance.

Importance of Historical Holiday Data (Variant 4: ST-Cross-Attn-No Holiday): This variant replaces the augmented time-series matrices with the standard time-series matrices of traffic flow, dropping off historical holiday traffic flow information. The results indicate a decrease in prediction accuracy, with performance worst at the 24-h prediction step. However, as the forecast horizon increases, prediction accuracy improves. This can be attributed to the specific traffic patterns on the first day of holidays, where traffic volume increases significantly. Without historical holiday features, the model struggles to capture this pattern, especially on the first day of the holiday (the first 24-h). However, in subsequent days, the traffic patterns become more stable and easier to predict. This validates the effectiveness of the augmented time-series matrix in enhancing holiday traffic flow prediction.

Finally, compared to the prediction performance in Table 1, although all ablation variants exhibit a decline in performance, they still outperform most baselines. This confirms that each module in the proposed model is essential and contributes to the overall enhancement of prediction accuracy.

4.8. Prediction Performance of Multi-Task Learning V.S. Single-Task Learning

This study employs a multi-task learning framework to jointly predict the inbound and outbound flow, allowing the sharing of spatial-temporal features between inbound and outbound flow, thereby enhancing the prediction accuracy. To evaluate the effectiveness of the multi-task learning approach, we compare its predictive accuracy with that of the single-task learning framework. Specifically, in addition to using the proposed multi-task learning model for the joint prediction of inbound and outbound flow, we employ a single-task learning framework to predict the inbound and outbound flow, separately. In the single-task learning framework, the attention interaction module is replaced with a standard attention mechanism.

The prediction results are presented in Table 6, demonstrating that, across various forecasting horizons, the multi-task learning model consistently outperforms the single-task learning approach for both inbound and outbound flow predictions. This confirms that multi-task learning enhances feature representation by enabling the sharing of spatial-temporal characteristics between inbound and outbound flow, thereby improving overall forecasting performance. Further analysis reveals a significant performance gap in outbound flow prediction between the multi-task and single-task learning approaches. This suggests that inbound flow provide valuable spatial-temporal information, which contributes to more accurate forecasting of outbound flow. In contrast, the prediction performance for inbound flow exhibits only a slight difference between the multi-task and single-task learning frameworks, suggesting an asymmetric interaction between sub-tasks. Specifically, the inbound flow prediction task significantly improves the accuracy of outbound flow prediction, whereas the outbound task contributes less to the inbound flow task. This asymmetry underscores the importance of explicitly modeling task dependencies when designing multi-task learning frameworks for spatiotemporal forecasting.

Table 6. Prediction result evaluation of multi-task learning (MTL) and single-task learning (STL).

	1 Day (24 Time Steps)			2 Days (48 Time Steps)			3 Days (72 Time Steps)		
	RMSE	MAE	WMAPE	RMSE	RMSE	MAE	WMAPE	MAE	RMSE
MTL (inbound flow)	136.89	66.91	28.55%	142.32	70.27	28.47%	148.34	72.08	27.47%
STL (inbound flow)	138.44	71.10	30.32%	144.04	73.90	29.94%	149.66	75.03	28.59%
MTL (outbound flow)	116.63	54.70	24.28%	122.86	59.54	25.37%	127.39	61.99	24.76%
STL (outbound flow)	139.93	74.74	31.87%	145.42	76.95	31.18%	152.67	80.72	30.78%

4.9. Analysis of Interactive Attention Between Inbound and Outbound Traffic Flows

This study introduces a novel cross-attention mechanism to capture the underlying evolutionary patterns between inbound and outbound flows, thereby enhancing their feature representations. To investigate its effectiveness, we visualize the temporal cross-attention between inbound and outbound flows during the holiday period. In the heatmap, (i, j) represents the attention weight (i.e., the correlation) between the inbound flow at time step i and the outbound flow at time step j . It is worth noting that the temporal cross-attention between inbound and outbound flows is symmetric. Therefore, we only visualize the cross-attention from inbound to outbound flows. The visualization is shown as Figure 12.

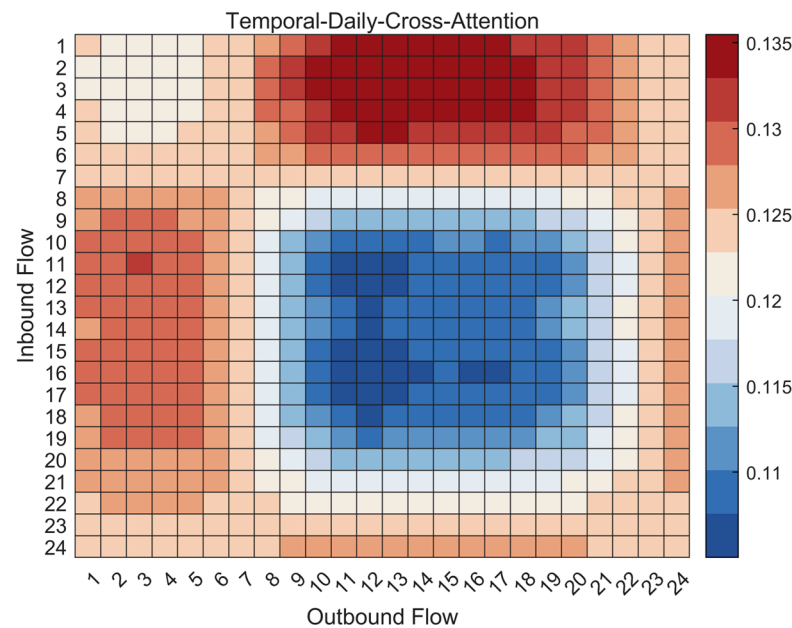


Figure 12. Heatmap of the temporal daily cross-attention Between inbound and outbound Flows.

As shown in the above figure, the temporal cross-attention between inbound and outbound flows exhibits a clear complementary pattern. High attention weights appear in regions where early inbound flows correspond to later outbound flows, and vice versa. This suggests that during the holiday period, the temporal evolution of inbound and outbound traffic flows tends to be opposite—inbound flow peaks earlier in the holiday period, while outbound flow tends to peak later. The model effectively captures this inverse temporal dependency, which aligns with typical passenger behavior patterns during holidays, such as outbound travel at the beginning and return travel at the end.

5. Conclusions

Highway traffic flow prediction during holidays is crucial for highway operation and management. Therefore, this study focuses on predicting the inbound and outbound flow at highway toll stations during holidays and proposes a joint prediction model based on multi-task learning. Specifically, the model takes the augmented time-series matrices of inbound and outbound flow as input to better capture the holiday characteristics. By leveraging the ST-Cross-Attn module that combines Bi-ConvLSTMs with Cross-Attn mechanism, the model effectively captures the complex spatial-temporal dependencies of inbound and outbound flow. Finally, a sequence-to-sequence framework is employed to jointly predict the future inbound and outbound flow over multi-time steps during holidays. Extensive experiments in real-world highway traffic flow datasets demonstrate its superior predictive performance and robustness in accurately forecasting highway traffic flow during holidays. Additionally, the ablation experiments confirm the significance of each model component and validate the effectiveness of the multi-task learning framework.

However, this study has some limitations. For instance, when predicting holiday traffic flow, the model currently relies solely on historical holiday traffic data. However, external factors such as weather conditions, policy restrictions, the popularity of tourist attractions, and other contextual information can also significantly influence traffic patterns during holidays. Future research could incorporate these external variables to better understand their impact on traffic evolution and further improve prediction accuracy. Moreover, due to limitations in data availability, this study evaluates the proposed model using highway traffic flow data during the Labor Day holiday. Experiments on traffic data from other holidays were not conducted. In future work, we aim to incorporate additional holiday

datasets—such as those from the New Year’s Day and National Day holidays—should such data become available, in order to further evaluate the model’s predictive performance across various holiday scenarios and to validate its robustness and generalizability.

Author Contributions: Conceptualization and writing—review and editing, J.Z. and S.Z.; methodology and investigation, X.L. and Y.Z.; software, Y.Z.; validation and visualization, Z.H.; formal analysis and writing—original draft preparation, H.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the project from National Engineering Research Center of System Technology for High-Speed Railway and Urban Rail Transit (2024YJ335).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is unavailable due to privacy or ethical restrictions.

Conflicts of Interest: Authors Xiaowei Liu, Yunfan Zhang, and Zhongyi Han are employed by the company Shandong Provincial Communications Planning and Design Institute Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Tian, J.; Zeng, J.; Ding, F.; Xu, J.; Jiang, Y.; Zhou, C.; Li, Y.; Wang, X. Highway traffic flow forecasting based on spatiotemporal relationship. *Chin. J. Eng.* **2024**, *46*, 1623–1629. [\[CrossRef\]](#)
2. Feng, T. Research on Prediction of Freeway Traffic Flow in Holidays Based on EMD and GS-SVM. Master’s Thesis, Chang’an University, Xi’an, China, 2016.
3. Zhang, J.; Mao, S.; Yang, L.; Ma, W.; Li, S.; Gao, Z. Physics-informed deep learning for traffic state estimation based on the traffic flow model and computational graph method. *Inf. Fusion* **2024**, *101*, 101971. [\[CrossRef\]](#)
4. Zhang, J.; Zhang, S.; Zhao, H.; Yang, Y.; Liang, M. Multi-frequency spatial-temporal graph neural network for short-term metro OD demand prediction during public health emergencies. *Transportation* **2025**, 1–23. [\[CrossRef\]](#)
5. Zhou, Z.; Huang, Q.; Wang, B.; Hou, J.; Yang, K.; Liang, Y.; Zheng, Y.; Wang, Y. Coms2t: A complementary spatiotemporal learning system for data-adaptive model evolution. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, 1–18. [\[CrossRef\]](#)
6. Ma, Z.; Xing, J.; Mesbah, M.; Ferreira, L. Predicting short-term bus passenger demand using a pattern hybrid approach. *Transp. Res. Part C Emerg. Technol.* **2014**, *39*, 148–163. [\[CrossRef\]](#)
7. Zhang, J.; Zheng, Y.; Qi, D. Deep spatio-temporal residual networks for citywide crowd flows prediction. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
8. Li, Y.; Yu, R.; Shahabi, C.; Liu, Y. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In Proceedings of the International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018.
9. Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Wang, P.; Lin, T.; Deng, M.; Li, H. T-GCN: A temporal graph convolutional network for traffic prediction. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 3848–3858. [\[CrossRef\]](#)
10. Zhang, J.; Chen, F.; Guo, Y.; Li, X. Multi-graph convolutional network for short-term passenger flow forecasting in urban rail transit. *IET Intell. Transp. Syst.* **2020**, *14*, 1210–1217. [\[CrossRef\]](#)
11. Wu, Z.; Pan, S.; Long, G.; Jiang, J.; Zhang, C. Graph wavenet for deep spatial-temporal graph modeling. In Proceedings of the 28th International Joint Conference on Artificial Intelligence, Macao, China, 10–16 August 2019; pp. 1907–1913.
12. Zhang, S.; Zhang, J.; Yang, L.; Wang, C.; Gao, Z. COV-STFormer for short-term passenger flow prediction during COVID-19 in urban rail transit systems. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 3793–3811. [\[CrossRef\]](#)
13. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017.
14. Jiang, J.; Han, C.; Zhao, W.; Wang, J. Pdformer: Propagation delay-aware dynamic long-range Transformer for traffic flow prediction. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 4365–4373.
15. Reza, S.; Ferreira, M.; Machado, J.; Tavares, J. A multi-head attention-based Transformer model for traffic flow forecasting with a comparative analysis to recurrent neural networks. *Expert Syst. Appl.* **2022**, *202*, 117275. [\[CrossRef\]](#)

16. Zhang, J.; Mao, S.; Zhang, S.; Yin, J.; Yang, L.; Gao, Z. EF-former for short-term passenger flow prediction during large-scale events in urban rail transit systems. *Inf. Fusion* **2025**, *117*, 102916. [\[CrossRef\]](#)
17. Yan, H.; Ma, X.; Pu, Z. Learning dynamic and hierarchical traffic spatiotemporal features with Transformer. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 22386–22399. [\[CrossRef\]](#)
18. Xu, M.; Dai, W.; Liu, C.; Gao, X.; Lin, W.; Qi, G.; Xiong, H. Spatial-temporal transformer networks for traffic flow forecasting. *arXiv* **2020**, arXiv:2001.02908.
19. Zhang, S.; Zhang, J.; Yang, L.; Chen, F.; Li, S.; Gao, Z. Physics guided deep learning-based model for short-term origin-destination demand prediction in urban rail transit systems under pandemic. *Engineering* **2024**, *41*, 276–296. [\[CrossRef\]](#)
20. Zang, D.; Ling, J.; Wei, Z.; Tang, K.; Cheng, J. Long-term traffic speed prediction based on multiscale spatio-temporal feature learning network. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 3700–3709. [\[CrossRef\]](#)
21. Qu, L.; Li, W.; Li, W.; Ma, D.; Wang, Y. Daily long-term traffic flow forecasting based on a deep neural network. *Expert Syst. Appl.* **2019**, *121*, 304–312. [\[CrossRef\]](#)
22. Wang, Z.; Su, X.; Ding, Z. Long-term traffic prediction based on lstm encoder-decoder architecture. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 6561–6571. [\[CrossRef\]](#)
23. Li, Y.; Chai, S.; Wang, G.; Zhang, X.; Qiu, J. Quantifying the uncertainty in long-term traffic prediction based on PI-ConvLSTM network. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 20429–20441. [\[CrossRef\]](#)
24. Zhang, S.; Zhang, J.; Yang, L.; Yin, J. Spatiotemporal attention fusion network for short-term passenger flow prediction on New Year's Day holiday in urban rail transit system. *IEEE Intell. Transp. Syst. Mag.* **2023**, *15*, 59–77. [\[CrossRef\]](#)
25. Qiu, H.; Zhang, J.; Yang, L.; Han, K. Spatial-temporal multi-task learning for short-term passenger inbound flow and outbound flow prediction on holidays in urban rail transit systems. *Transportation* **2025**, 1–30.
26. Wang, A.; Ye, Y.; Song, X.; Zhang, S.; Yu, J.J. Traffic prediction with missing data: A multi-task learning approach. *IEEE Trans. Intell. Transp. Syst.* **2023**, *24*, 4189–4202. [\[CrossRef\]](#)
27. Zhang, K.; Liu, Z.; Zheng, L. Short-term prediction of passenger demand in multi-zone level: Temporal convolutional neural network with multi-task learning. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 1480–1490. [\[CrossRef\]](#)
28. Yi, Z.; Zhou, Z.; Huang, Q.; Chen, Y.; Yu, L.; Wang, X.; Wang, Y. Get Rid of Isolation: A Continuous Multi-task Spatio-Temporal Learning Framework. In Proceedings of the Thirty-Eighth Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 10–15 December 2024.
29. Yang, Y.; Zhang, J.; Yang, L.; Gao, Z. Network-wide short-term inflow prediction of the multi-traffic modes system: An adaptive multi-graph convolution and attention mechanism based multitask-learning model. *Transp. Res. Part C Emerg. Technol.* **2024**, *158*, 104428. [\[CrossRef\]](#)
30. Zou, G.; Lai, Z.; Wang, T.; Liu, Z.; Li, Y. Mt-stnet: A novel multi-task spatiotemporal network for highway traffic flow prediction. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 8221–8236. [\[CrossRef\]](#)
31. Azad, R.; Asadi-Aghbolaghi, M.; Fathy, M.; Escalera, S. Bi-directional ConvLSTM U-Net with densely connected convolutions. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Seoul, Republic of Korea, 27–28 October 2019.
32. Lin, Z.; Li, M.; Zheng, Z.; Cheng, Y.; Yuan, C. Self-attention ConvLSTM for spatiotemporal prediction. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34.
33. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.; Wong, W.; Woo, W. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; Volume 18.
34. Liu, L.; Zhu, Y.; Li, G.; Wu, Z. Online metro origin-destination prediction via heterogeneous information aggregation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 3574–3589. [\[CrossRef\]](#)
35. Sapankevych, N.; Sankar, R. Time series prediction using support vector machines: A survey. *IEEE Comput. Intell. Mag.* **2009**, *4*, 24–38. [\[CrossRef\]](#)
36. Zhang, Y.; Shi, X.; Zhang, S.; Abraham, A. A XGBoost-based lane change prediction on time series data using feature engineering for autopilot vehicles. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 19187–19200. [\[CrossRef\]](#)
37. Wang, L.; Zeng, Y.; Chen, T. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Syst. Appl.* **2015**, *42*, 855–863. [\[CrossRef\]](#)
38. Wang, K.; Li, K.; Zhou, L.; Hu, Y. Multiple convolutional neural networks for multivariate time series prediction. *Neurocomputing* **2019**, *360*, 107–119. [\[CrossRef\]](#)

39. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The performance of LSTM and BiLSTM in forecasting time series. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019.
40. Yamak, P.; Li, Y.; Gadosey, P. A comparison between ARIMA, LSTM, and GRU for time series forecasting. In Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence, Sanya, China, 20–22 December 2019.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.