

基于GCN-Transformer模型的短时交通量确定性预测

廖梦媛¹, 戎 荣¹, 张梦琦¹, 左安洋¹, 莫乾群²

¹重庆交通大学信息科学与工程学院, 重庆

²重庆交通大学数学与统计学院, 重庆

收稿日期: 2024年4月24日; 录用日期: 2024年7月10日; 发布日期: 2024年7月19日

摘要

为提高短时交通量的预测精度, 充分利用交通网络的拓扑结构与时序信息, 本文提出了一种基于GCN-Transformer模型的短时交通量确定性预测。首先读取并预处理四个方向的交通流量数据, 构建路网结构邻接矩阵, 划分数据集。然后利用GCN提取特征并进行图卷积操作, 将输出作为Transformer模型的输入来学习时序信息。最后使用训练好的模型进行预测, 并计算评价指标。研究结果表明, 本文所提出的混合预测模型GCN-Transformer拥有较好的结果, 在采用的四个评价指标中, RMSE、MAE、RMSPE和MAPE的值分别为9.886%、7.977%、11.018%、11.734%。

关键词

GCN-Transformer模型, 短时交通量, 确定性预测, 预测精度

Deterministic Prediction of Short-Term Traffic Volume Based on GCN-Transformer Model

Mengyuan Liao¹, Rong Rong¹, Mengqi Zhang¹, Anyang Zuo¹, Qianqun Mo²

¹School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing

²School of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing

Received: Apr. 24th, 2024; accepted: Jul. 10th, 2024; published: Jul. 19th, 2024

Abstract

In order to improve the prediction accuracy of short-term traffic volume and make full use of the

文章引用: 廖梦媛, 戎荣, 张梦琦, 左安洋, 莫乾群. 基于 GCN-Transformer 模型的短时交通量确定性预测[J]. 交通技术, 2024, 13(4): 234-243. DOI: 10.12677/ojtt.2024.134027

topology and timing information of traffic network, a deterministic prediction of short-term traffic volume based on GCN-Transformer model is proposed in this paper. Firstly, the traffic flow data in four directions are read and preprocessed, the network structure adjacency matrix is constructed, and the data set is divided. Then, GCN is used to extract features and perform graph convolution operations, and the output is used as the input of Transformer model to learn timing information. Finally, the trained model is used to predict and the evaluation index is calculated. The research results show that the hybrid prediction model GCN-Transformer proposed in this paper has better results, and the values of RMSE, MAE, RMSPE and MAPE among the four evaluation indicators are 9.886%, 7.977%, 11.018% and 11.734%, respectively.

Keywords

GCN-Transformer Model, Short-Term Traffic Volume, Deterministic Prediction, Prediction Accuracy

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着城市人口的不断增加，机动车数量快速增长，公路路网密度不断增大[1]。这种情况导致城市交通系统面临巨大压力，出现了环境污染、交通拥堵等问题，进而导致道路利用率低下、道路环境污染、能源浪费和交通事故的发生。交通流量预测分为短期和中长期两类，短时交通流量通常是指 5 到 15 分钟的流量预测。交通量的预测本身是一个时间序列的问题[2] [3]，在不确定性条件和缺乏数据资料的情况下，交通量的预测是十分复杂的。

目前，世界各地都在不断增加对智能交通系统的投入，许多的研究也就围绕此展开。2016 年，罗来鹏[4]利用 BP 神经网络方法预测了相关内容。2016 年，Yuan 等[5]结合 KNN 算法确定了最优聚类数 K，选择 BP 神经网络对短时交通流进行了预测。2017 年，罗文慧等[6]提出了一种基于深度学习的短时交通流预测模型。2018 年，Tian 等[7]提出了一种基于长短期记忆(LSTM)的新方法。同年，罗向龙等[8]提出了一种基于 K-最近邻(KNN)与长短时记忆(LSTM)网络模型相结合的短时交通流预测模型。2020 年，张珊[9]设计了两种不同的循环神经网络(RNN): 长短期记忆(LSTM)网络和门控制循环单元(GRU)网络。2023 年，高榕等[10]提出了一种基于编解码器改进的时空 Transformer 模型(ISTTM)。

目前的交通流预测研究虽然已经取得了一些进展，但由于交通具有复杂性和受多种因素(如天气)的影响，现有方法在实际预测需求方面还有待优化，因此本文建立了 GCN-Transformer 来对短时交通量进行确定性预测。

2. 基于 GCN-Transformer 的确定性预测研究

2.1. 对比模型的预测原理

2.1.1. GCN 模型

图卷积神经网络(Graph Convolutional Network, GCN)是一种基于图数据的深度学习方法，它利用图结构中节点和边之间的拓扑结构，处理图像标签之间的空间关系[11]。现有的图卷积神经网络分为谱方法和空间方法两类，谱方法利用图卷积定理从谱域定义图卷积，而空间方法从节点域出发，通过定义聚合函

数来聚合每个中心节点和其邻近节点。而本文主要采用了空间方法, 将某段道路的交通流数据通过 GCN, 将数据通过点与边来表示, 转变为网络拓扑结构(图 1)。

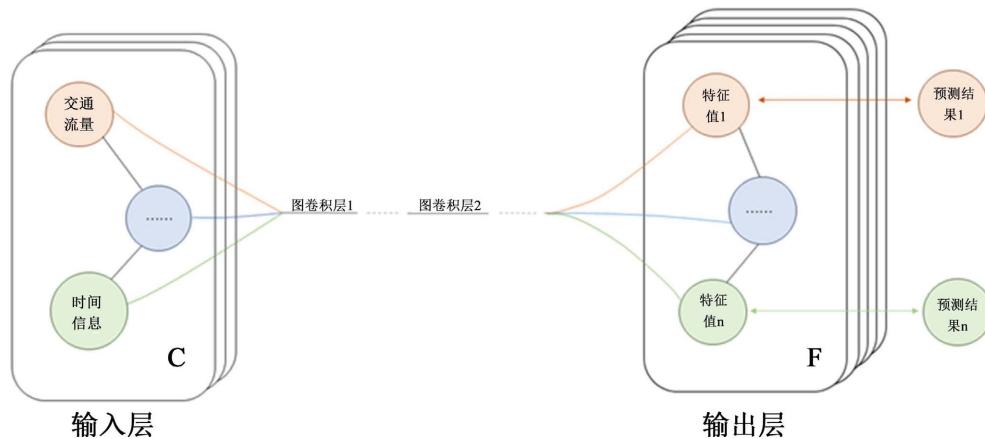


Figure 1. Schematic diagram of GCN model network structure

图 1. GCN 模型网络结构示意图

图卷积神经网络主要包括卷积算子和池化算子的构建, 其中卷积算子的目的是刻画节点的局部结构, 而池化算子的目的是学到网络的层级化表示, 降低参数[12]。在解决节点级别的任务时, 研究人员更关注如何给每个节点学到更好的表达, 此时池化算子并不必要, 因此前期大量的工作仅关注图上卷积算子的构建, 而池化算子通常主要应用在图级别的任务上。

计算公式如下:

$$H^{l+1} = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right) \quad (1)$$

其中, $H^{(l)}$ 表示第 l 层的节点表示, N 是图中节点的数量, d 是节点的特征维度。 $\tilde{A} = A + I_N$ 是图邻接矩阵 A 与自环矩阵 I_N 的和, I_N 是一个 $N \times N$ 的单位矩阵。 \tilde{D} 是对角矩阵, 其中 $D_{ii} = \sum_j \tilde{A}_{ij}$ 。 $W^{(l)} \in R^{d \times d'}$ 是第 l 层的权重矩阵, d' 表示第 $l+1$ 层的节点特征维度。 $\sigma(\cdot)$ 表示激活函数, 常用的激活函数有 ReLU、sigmoid 等。

GCN 的优势在于能够利用图结构中的局部与全局信息, 从而更好地学习节点的表示。通过聚合邻居节点的特征, GCN 能够传播信息到远离节点, 捕捉到跨节点的联系和依赖关系[13]。而传统方法只能考虑节点的局部邻居信息, 无法获取整体的图结构特征。

2.1.2. Transformer 模型

交换器(Attention Is All You Need, Transformer)是一种用于序列建模的深度学习模型, 最初由 Vaswani 等人于 2017 年提出, 并在机器翻译任务中取得了重大突破。与传统的循环神经网络(Recurrent Neural Network, RNN)不同, Transformer 采用了自注意力机制, 通过并行计算来处理序列中的不同位置, 在多个层次上进行信息交换和集成, 从而提高了模型的学习能力和效果(图 2)。

Transformer 模型的核心结构包括编码器和解码器。编码器负责将输入序列进行编码, 而解码器则生成输出序列[14]。每个编码器和解码器层都由多个子层组成, 其中包括自注意力机制和前馈神经网络。

自注意力机制是 Transformer 模型的关键部分, 它能够捕捉序列中不同位置之间的依赖关系[15]。其计算公式如下:

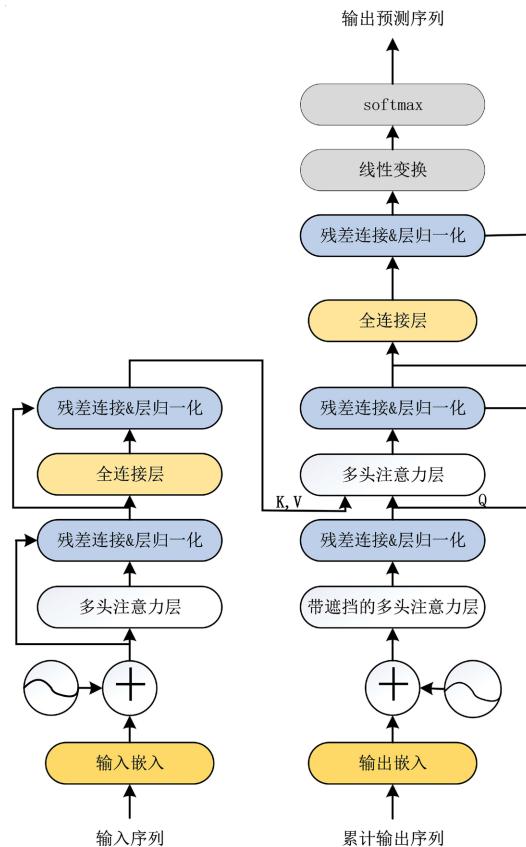


Figure 2. Schematic diagram of Transformer model network structure
图 2. Transformer 模型网络结构示意图

$$\alpha_{ij} = \text{softmax} \left(\frac{Q(x_i) K(x_j)^T}{\sqrt{d_K}} \right) V \quad (2)$$

在上述公式中， Q 、 K 和 V 分别表示线性变换得到的查询(query)、键(key)和值(value)向量。 d_K 是特征维度的平方根，用于缩放相似度。通过自注意力机制，模型能够根据输入序列在不同位置之间建立起动态的依赖关系。

除了自注意力机制，Transformer 模型还引入了位置编码来表示序列中每个位置的相对位置信息。位置编码是一个与特征维度相同的向量序列，在计算过程中被添加到输入特征中。通常，位置编码可以通过正弦和余弦函数来计算，具体公式如下：

$$PE_{(pos,2i)} = \sin \left(pos / 10000^{2i/d_{\text{model}}} \right) \quad (3)$$

$$PE_{(pos,2i+1)} = \cos \left(pos / 10000^{2i/d_{\text{model}}} \right) \quad (4)$$

在上述公式中， pos 表示位置， i 表示位置编码的维度， d_{model} 表示特征维度。通过位置编码，模型能够感知输入序列中不同位置的信息。

总结起来，Transformer 模型是一种强大的深度学习模型，适用于各种序列建模任务。它通过自注意力机制和位置编码，能够捕捉序列中的全局依赖关系，并实现并行计算[16]，提高了模型的效率和学习能力。

2.1.3. Bi-LSTM 模型

长短时记忆网络(Long Short-Term Memory, LSTM)是一种循环神经网络结构的变体，能够解决循环神经网络中遇到的梯度爆炸和梯度消失问题[17]。遗忘门、输入门和输出门这三种特殊的结构在 LSTM 体系中起到了关键作用，使得 LSTM 能够很好地处理时间序列数据中的短期和长期时间序列内的相关性[18]。LSTM 数学原理运用公式如式(5)~(10)所示：

$$f_t = \sigma(W_f [h_{t-1}, x_t] + b_f) \quad (5)$$

$$i_t = \sigma(W_i [h_{t-1}, x_t] + b_i) \quad (6)$$

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \quad (7)$$

$$\bar{C}_t \tanh(W_c [h_{t-1}, x_t] + b_c) \quad (8)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tanh(W_c [h_{t-1}, x_t] + b_c) \quad (9)$$

$$h_t = o_t \times \tanh(C_t) \quad (10)$$

其中， W_f, W_i, W_o, W_c 分别表示遗忘门、输入门、输出门和记忆单元的权重系数矩阵； \tanh 函数表示双曲正切激活函数； σ 表示 Sigmoid 函数； b_f, b_i, b_o, b_c 表示偏置向量。

2.2. 混合预测模型(GCN-Transformer)构建流程

2.2.1. 数据预处理阶段

我们首先对四个方向的交通流量数据进行读取与预处理，这些方向包括南北和东西两个主要流向。接下来，我们将这些不同方向的交通流量数据融合，以便构建一个统一的交通流量数据表示。在此基础上，我们进一步构建了路网结构的邻接矩阵，该矩阵能够有效刻画路网中不同节点之间的连接关系。最后，我们根据研究需求，将数据划分为训练集、验证集和测试集，为后续模型训练与评估提供数据支撑。

2.2.2. 模型训练阶段

在模型训练阶段，我们利用图卷积网络(GCN)对合并后的交通流量数据进行特征提取和图卷积操作。GCN 的引入使我们能够充分捕捉路网中复杂的拓扑结构和空间依赖关系。随后，我们将 GCN 的输出作为 Transformer 模型的输入，利用 Transformer 强大的时序建模能力，进一步学习交通流量数据中的时序信息。在模型训练过程中，我们采用适当的损失函数，并通过反向传播算法对模型参数进行优化，以实现模型性能的最大化。

2.2.3. 模型评估阶段

完成模型训练后，我们使用训练好的模型对测试集数据进行预测，并计算一系列评价指标，包括均方根误差(RMSE)、平均绝对误差(MAE)、均方根百分比误差(RMSPE)和平均百分比误差(MAPE)。这些指标能够全面反映模型在交通流量预测任务上的性能表现，为后续的研究与应用提供重要参考。

2.3. 模型预测精度评价指标

确定性预测是指在给定输入条件下，通过建立预测模型来预测未来事件或结果的过程。确定性预测假定未来的结果是可预测和确定的，即未来的结果只有一个可能的发生方式。

本文的短时交通流，便是基于已知的某路段的数据信息和可观察的因素，通过分析、数据预处理以及建立相关的 GCN-Transformer 模型来估计未来的短时间内交通流量的数据。

因此，本文主要采用以下评价指标：

1) 平均绝对误差(Mean Absolute Error, MAE):

$$\text{MAE} = \frac{1}{n} \times \sum_{i=1}^n |y_i - y'_i| \quad (11)$$

其中, n 表示样本数, y_i 表示第 i 个样本的真实值, y'_i 表示第 i 个样本的预测值。

2) 均方根误差(Root Mean Squared Error, RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (y_i - y'_i)^2} \quad (12)$$

其中, n 表示样本数, y_i 表示第 i 个样本的真实值, y'_i 表示第 i 个样本的预测值。

3) 平均绝对百分比误差(Mean Absolute Percentage Error, MAPE):

$$\text{MAPE} = \frac{1}{n} \times \sum_{i=1}^n \left| \frac{(y_i - y'_i)}{y_i} \right| \quad (13)$$

其中, n 表示样本数, y_i 表示第 i 个样本的真实值, y'_i 表示第 i 个样本的预测值。

4) 相对均方根误差(Root Mean Squared Error Relative, RMSRE):

$$\text{RMSRE} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\bar{y}} \quad (14)$$

其中, n 表示样本数, y_i 表示第 i 个样本的真实值, y'_i 表示第 i 个样本的预测值, \bar{y} 表示均值。

通过使用这些评价指标, 可以对预测模型的准确性进行客观的量化评估。

3. 应用案例和实际验证

3.1. 交通量原始数据描述及预处理

3.1.1. 数据来源

为了训练模型, 使其拥有可行性, 我们对真实的现场交通数据进行了调研。通过来自重庆主城区某主干道的交叉口收集到的数据进行训练。收集时间为一星期, 统计间隔为 5 min 作为样本数据的数据集。将生成的数据集按照 3:1:1 的比例划分出训练集、验证集、测试集。

3.1.2. 数据预处理

交通流量数据通常由多个传感器采集。由于一些不可控因素, 例如传感器故障或者天气等, 数据中可能会出现缺失值。而缺失值的存在会影响到基于数据的交通流量预测的准确性和稳定性, 因此需要进行处理以保证预测的可靠性。

在本文中, 我们采用将缺失值使用均值填充的方法。这种方法基于一个简单的假设: 使用该列中的均值来替换缺失值, 利用已有的数据来估计缺失数据的取值。具体的处理步骤如下:

1) 计算均值:

针对车流量这一特征, 通过计算其对应列的均值来获取当前数据集中的平均车流量值。

2) 缺失数据填充:

对于每个缺失的数据点, 采用车流量列的均值来进行填充, 确保数据的完整性和连续性。

3) 数据准备:

将填充后的数据转换成二维向量, 并准备用于模型训练或预测。

构建了路网结构的邻接矩阵, 为后续模型训练提供了数据基础。

采用均值填充缺失值的优点在于简单易行，且能够较好地保留各特征的整体分布和趋势。同时，均值填充也可以避免出现 NAN 或者 Infinity 等无穷大的值，使得数据更为完整。然而，均值填充也有一些缺点，例如对于随机分布的数据可能会导致部分的数据抖动，且对于数据的分布不稳定情况可能会引入偏差等问题。因此，应用中需根据具体情况选择合适的数据预处理方式。

在短时交通流量的确定性预测中，数据预处理是一个至关重要的阶段。填充缺失值使用均值填充的方法在简化问题的同时，也能够提高数据利用率，保证数据的完整性和可靠性。但是，需要注意的是选择预处理方法时要根据具体任务和数据特征进行选择，以保证最终的预测结果准确可靠。

3.2. 案例结果与分析

通过文件和代码来建立网络拓扑结构，在获取到测试数据与预测数据后，对预测模型的预测性能展开评估。算出模型和其他相关模型的预测精度评价指标数值，涵盖均方根误差、平均绝对误差、均方根相对误差、平均绝对百分比误差。这些指标在预测模型的评估中常被用来比较不同模型的性能。具体计算结果展示于表 1 中。

Table 1. Evaluation index values of different prediction models

表 1. 不同预测模型的评价指标值

	RMSE	MAE	RMSPE	MAPE
GCN	23.246	13.466	18.519	20.417
Transformer	31.587	28.694	57.951	38.346
Bi-LSTM	39.562	37.729	36.739	25.449
GCN-Transformer	9.886	7.977	11.018	11.734

实验结果表明，GCN-Transformer 在各项指标上均表现出色，尤其在 RMSE 和 MAE 方面，显著优于其他模型。这表明，结合图卷积网络与 Transformer 的混合模型在处理时序预测任务中具有较强的优势。

下面将采用预测结果对比折线图来详细描绘交通量随时间变化的动态趋势。图中，纵轴精确反映了交通量的数值变化，而横轴则代表以每日零点为起点的时间轴，每个时间步长精准设定为 5 分钟。

GCN 模型的预测结果如下图 3 所示。

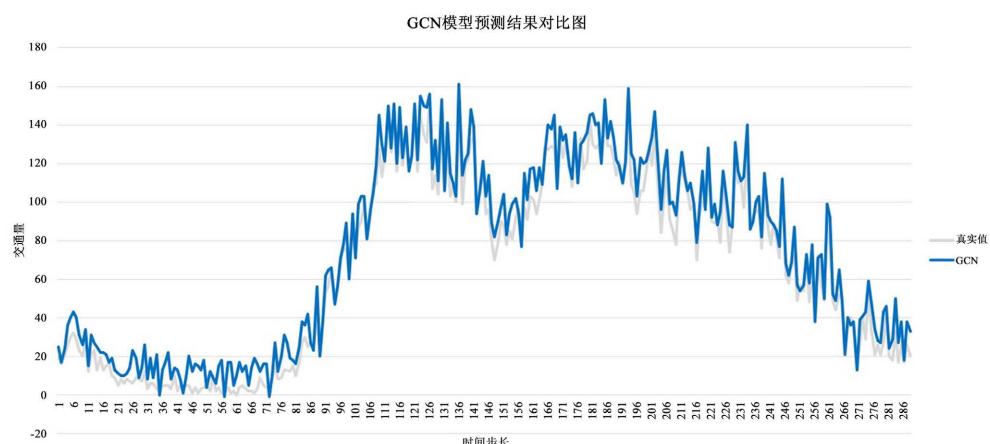


Figure 3. Comparison of GCN model prediction results

图 3. GCN 模型预测结果对比图

Transformer 模型的预测结果如下图 4 所示。

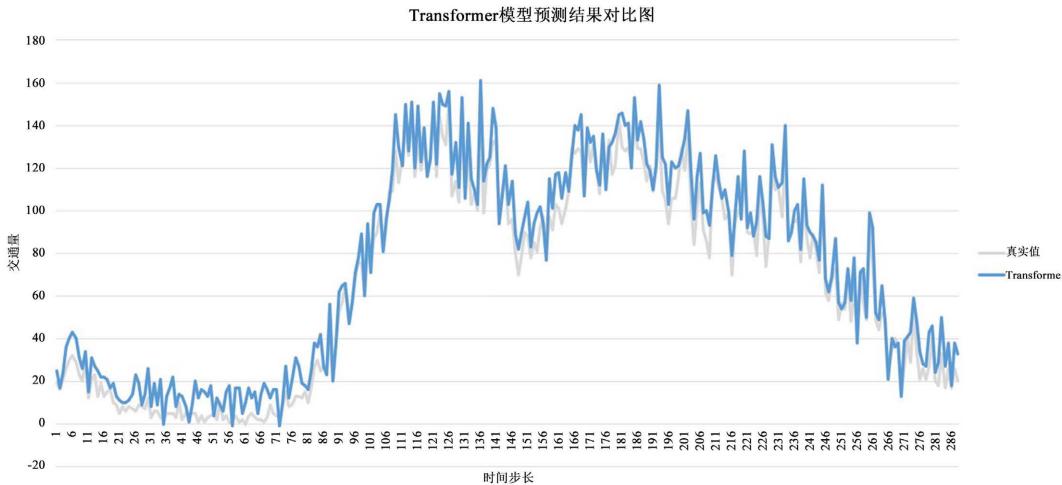


Figure 4. Comparison of Transformer model prediction results

图 4. Transformer 模型预测结果对比图

Bi-LSTM 模型的预测结果如下图 5 所示。

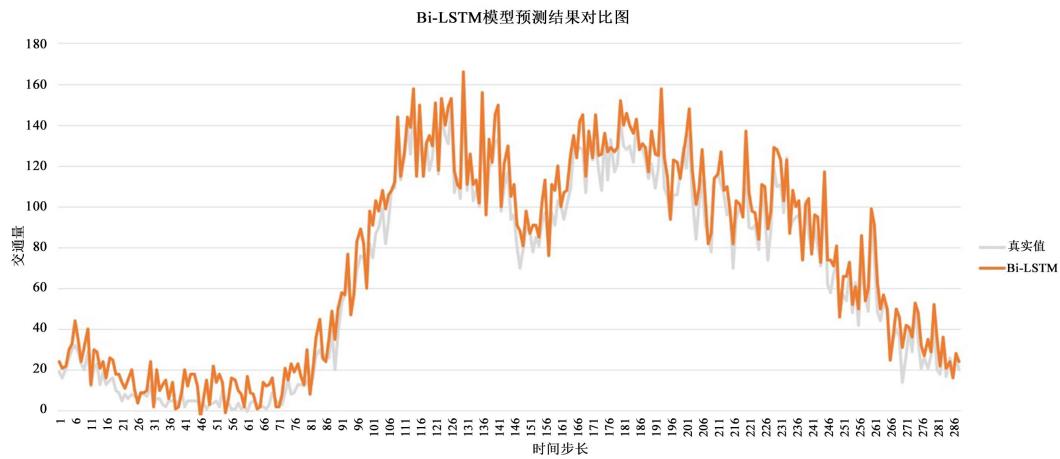


Figure 5. Comparison of Bi-LSTM model prediction results

图 5. Bi-LSTM 模型预测结果对比图

GCN-Transformer 模型的预测结果如下图 6 所示。

3.3. 预测结果分析

通过分析表 1、图 4~6，可以得到：

本文所提出的 GCN-Transformer 混合预测模型在短时交通量确定性预测中拥有较好的结果，在采用的四个评价指标中的值分别为 RMSE = 9.886%、MAE = 7.977%、RMSPE = 11.018%、MAPE = 11.734%，从图 6 中更是可以直观地看到 GCN-Transformer 混合预测模型的预测值与真实值重合度较高。

与效果较好的对比模型 GCN 相比，本文所提模型的四项评价指标 RMSE、MAE、RMSPE 和 MAPE 分别改进了 13.36%、5.489%、7.501% 和 8.683%。这说明因为混合模型融合了 GCN 和 Transformer 它们

的优点，从而提高了预测精度，而且混合模型能够综合利用多种信息，提供更全面的特征表示，适当增加了模型的复杂度，提高了对数据的拟合能力。

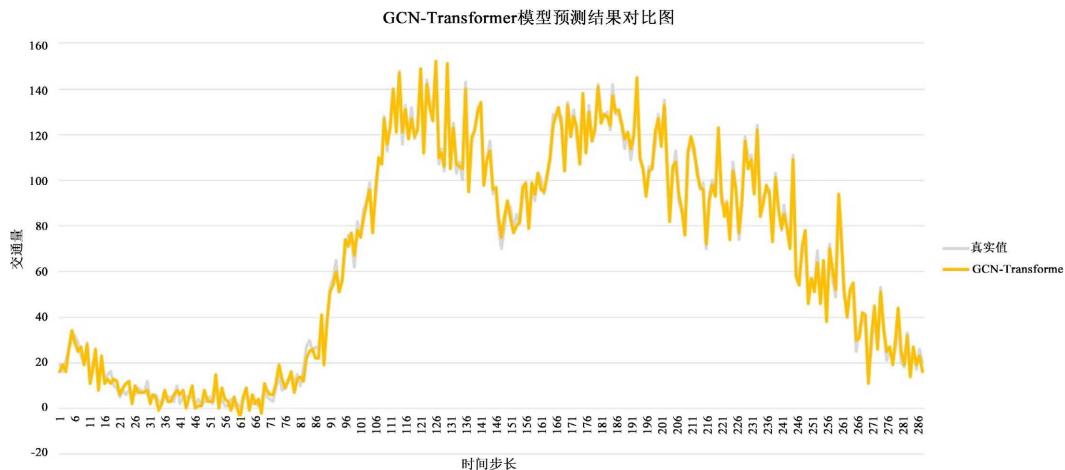


Figure 6. Comparison of GCN-Transformer model prediction results
图 6. GCN-Transformer 模型预测结果对比图

4. 主要内容和结论

在本文中，我们提出了 GCN-Transformer 组合模型的交通流量预测方法。GCN-Transformer 模型结合了图卷积神经网络和自注意力机制，能够处理交通网络中的复杂拓扑结构和时空关系。通过图卷积神经网络，模型可以有效地捕捉节点之间的相互作用，而自注意力机制则有助于对时间和空间维度的信息进行有效的编码和建模。结合图神经网络挖掘非欧数据时空特征的优点和自注意力机制的特点，对输入数据进行线性转换后分别提取时间特征和空间特征，通过多次跳层连接进行加权融合，最后经过线性转换输出预测数据。结果表明，GCN-Transformer 组合模型对数据的拟合效果良好，具有较高的预测精度，预测的交通流数据保持了较高的一致性，是一种实用的短时交通流预测模型，满足了最开始我们希望通过该组合模型对交通流进行预测，减少交通拥堵等问题的设想。

通过本文的研究，我们得出以下结论：GCN-Transformer 混合模型在短时交通量确定性预测中提高了预测精度和准确性，对交通流数据的预测效果明显优于单一模型 GCN、Transformer 和混合模型 Bi-LSTM，为交通流预测提供了一种有效的方法。

综合以上分析，GCN-Transformer 组合模型满足了本文对交通流预测的实际需求，具有较高的预测精度和实用性。该模型为解决交通拥堵等问题提供了有效的方法和工具，为交通管理和规划提供了有力支持。在未来的研究中，可以进一步优化模型结构和算法，提升预测效果，同时扩展模型在其他领域的应用，以实现更广泛的社会价值。

致 谢

在本研究完成之际，我们要衷心感谢学校对我们研究工作的大力支持。同时，也要向所有给予我们帮助和指导的老师表达深深的谢意，是你们的支持与鼓励，让我们的研究得以顺利进行。最后，感谢所有项目组的成员，你们的存在让此次研究之旅更加充实和有意义。

基金项目

全国大学生创新创业训练项目，项目编号：S202310618001。

参考文献

- [1] 缪荣辉. 考虑突发事件影响下的路网交通量预测方法研究[J]. 交通科技与管理, 2023(17): 27-29.
- [2] 丁志坤, 朱梦炼, 宋义勇. 基于改进“四阶段法”的高速公路交通量预测研究[J]. 重庆交通大学学报(自然科学版), 2017, 36(5): 86-90.
- [3] 刘宗明, 贾志绚, 李兴莉. 基于灰色马尔科夫链模型的交通量预测[J]. 华东交通大学学报. 2012, 29(1): 30-34.
- [4] 罗来鹏. 基于 BP 神经网络的城市占道交通拥堵预测[J]. 黑龙江工程学院学报, 2016, 30(1): 48-50.
- [5] Yuan, Z.W., Zhang, W.W. and Yang, M. (2016) A Short-Term Traffic Flow Prediction Approach of Neural Network Based on Cluster Analysis. *Proceedings of 2016 3rd International Conference on Engineering Technology and Application*, Phuket, 14-15 January 2016, 152-157.
- [6] 罗文慧, 董宝田, 王泽胜. 基于 CNN-SVR 混合深度学习模型的短时交通流预测[J]. 交通运输系统工程与信息, 2017, 17(5): 68-74.
- [7] Tian, Y., Zhang, K.L., Li, J.Y., Lin, X.X. and Yang, B.L. (2018) LSTM-Based Traffic Flow Prediction with Missing Data. *Neurocomputing*, 318, 297-305. <https://doi.org/10.1016/j.neucom.2018.08.067>
- [8] 罗向龙, 李丹阳, 杨彧, 张生瑞. 基于 KNN-LSTM 的短时交通流预测[J]. 北京工业大学学报, 2018, 44(12): 1521-1527.
- [9] 张珊. 基于 LSTM 和 GRU 熵权集成的短时交通流预测研究[D]: [硕士学位论文]. 青岛: 山东科技大学, 2019.
- [10] 高榕, 万以亮, 邵雄凯, 吴歆韵. 面向改进的时空 Transformer 的交通流量预测模型[J]. 计算机工程与应用, 2023, 59(7): 250-260.
- [11] 王雨松, 吴向东, 尤晨欣, 等. 基于 DWT-GCN 的短时交通流预测[J]. 计算机系统应用, 2022, 31(9): 306-312.
- [12] 徐冰冰, 岑科廷, 黄俊杰, 等. 图卷积神经网络综述[J]. 计算机学报, 2020, 43(5): 755-780.
- [13] 史昕, 曹凤腾, 纪艺, 等. 基于多尺度时空特征和软注意力机制的交通流预测方法[J]. 计算机工程, 2024, 4(24): 1-13.
- [14] Li, G., Zhong, S., Xiang, L., et al. (2021) A Lightweight and Accurate Spatial-Temporal Transformer for Traffic Forecasting. *IEEE Transactions on Knowledge and Data Engineering*, 35, 10967-10980. <https://doi.org/10.1109/TKDE.2022.3233086>
- [15] 张力. 基于 Transformer 的短时交通流预测[J]. 信息与电脑(理论版), 2022, 34(7): 66-68.
- [16] 杨国亮, 习浩, 龚家仁, 等. 基于 Transformer 的短时交通流时空预测[J]. 计算机应用与软件, 2024, 41(3): 169-173+225.
- [17] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, 9, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [18] 张阳, 胡月, 陈德旺, 等. 基于 GCN-BiLSTM 的短时交通流预测模型[J]. 武汉理工大学学报(交通科学与工程版), 2023, 47(5): 802-806.