

文章编号: 1674-599X(2020)03-0074-07

## 基于深度学习的短时交通流预测模型

卢生巧, 黄中祥

(长沙理工大学 交通运输工程学院, 湖南 长沙 410004)

**摘要:** 针对现有交通流短时预测模型在描述交通流时空特征能力较弱的问题, 提出了一种卷积神经网络和门控循环单元神经网络相结合的深度神经网络预测模型。该模型利用卷积神经网络提取短时交通流数据的空间特征, 并将结果输入到门控循环单元神经网络中, 挖掘短时交通流数据的时间特征。以加州交通局绩效评估系统的交通流数据为例, 对该模型进行训练, 验证该模型的准确性。试验结果表明: 与现有的模型相比, 该模型具有更好的预测性能, 其平均绝对百分误差显著减少, 可为交通管理与控制提供有效依据。

**关键词:** 短时预测; 深度学习; 卷积神经网络; 门控循环单元神经网络

中图分类号: U491.14 文献标志码: A

DOI:10.16544/j.cnki.cn43-1494/u.2020.03.012

### Prediction model of short-term traffic flow based on CNN-GRU deep learning

LU Sheng-qiao, HUANG Zhong-xiang

(College of Transportation Engineering, Changsha University of Science & Technology, Changsha 410004, China)

**Abstract:** To improve the prediction performance of temporal-spatial characteristics by the existing short-term traffic flow prediction model, a deep neural network prediction model combining convolutional neural network (CNN) and gated recurrent unit (GRU) neural network is proposed. In the model, the CNN is firstly used to extract the spatial characteristics of short-term traffic flow data, and then the results are input into the GRU to mine the temporal features. Taking the traffic flow data in the California Department of Transportation's performance evaluation system as an example, the model is trained to verify its accuracy. The results show that compared with existing models, the proposed prediction model has better prediction performance. The average absolute percentage error is significantly reduced, which can provide an effective basis for traffic management and control.

**Key words:** short-term prediction; deep learning; convolutional neural network; gated recurrent unit

交通流预测是城市智能交通系统(Intelligent Transport System, 简称为 ITS)的重要组成部分。精确的短时交通流预测, 可为道路使用者提供实时的、有价值的参考信息, 帮助其选择更优的出行路径, 也可预测道路拥堵的地点和时间。运输从业者可以提前选择优化路径, 减少道路拥堵<sup>[1]</sup>。短时交通流预测有助于 ITS 的道路交通控制和交通诱导, 提高道路安全性<sup>[2-3]</sup>。

目前, 短时交通流预测方法有基于模型算法和

数据驱动的方法。其中, 基于模型算法包括: 卡尔曼滤波器<sup>[4]</sup>、梯度增强树回归、时间序列的自回归统计及贝叶斯网络等经典的浅层学习算法。Hamed<sup>[5]</sup>等人利用 Box-Jenkins 技术, 建立 ARIMA 模型, 实现了交通量的短期预测。张涛<sup>[6]</sup>等人应用基于 K 近邻的非参数回归预测模型, 利用 K 值构造的预测区间, 进行特殊路况的预测, 获得了较高的预测精度。Hu<sup>[7]</sup>等人利用粒子群算法, 优化支持向量机(Support Vector Machine, 简称为 SVR)参数,

收稿日期: 2019-12-02

基金项目: 国家自然科学基金(51078044, 51338002)

作者简介: 卢生巧(1994-), 男, 长沙理工大学硕士生。

进行预测。

这些浅层结构的预测模型, 处理小样本数据时, 简单高效。处理大样本数据和复杂非线性函数时, 存在局限性。深度学习和神经网络的组合不断应用于交通流短期预测的深层结构模型中, 推动了基于数据驱动方法的神经网络和深度学习<sup>[8-9]</sup>等迅速发展, 提供了更准确的交通流预测工具。其中, 循环神经网络(Recurrent Neural Networks, 简称为 RNN)为变体的长短时记忆神经网络(Long Short-Term Memory, 简称为 LSTM), 可提取数据的长期依赖特征<sup>[10]</sup>。Polson<sup>[11]</sup>等人使用 Math Container 方法, 结合 tanh 层的深度学习模型, 成功预测特殊条件下急剧变化的交通量。Yu<sup>[12]</sup>等人利用深度卷积神经网络和长短时记忆网络相结合的深度结构, 捕获交通网络路径的时空相关性, 验证了卷积神经网络 CNN(Convolutional Neural Network, 简称为 CNN)和 LSTM 结合的可行性。刘明宇<sup>[13]</sup>等利用门控循环单元神经网络 (Gated Recurrent Unit, 简称为 GRU)进行短时交通流预测, 取得较高的预测精度。作者基于时空关联矩阵, 拟利用深度神经网络挖掘短时交通流数据的时空特征。采用 Python 软件, 将 CNN 和 GRU 两者结合起来, 构建一种混合神经网络预测模型, 并以加州交通局绩效评估系统 PeMS(Performance Measurement System, 简称为 PeMS)的实例数据为依托, 对该模型进行训练和预测。

## 1 基于 CNN-GRU 的短时交通流预测模型

交通流在空间域和时间域上具有明显的特征。在空间域中, 某些位置上的交通流, 可能与附近地区有更强的依赖性(拓扑局部性)。在时间域中, 部分历史交通流, 会对未来的交通流产生长期影响。

### 1.1 CNN 模型结构

根据经验和实验数据, 模型选择 CNN 层作为数据输入层, 预测效果更好。因此, 选择使用一维 CNN 结构, 挖掘交通流数据中的空间特征。CNN 是一种深度神经网络, 能够捕获交通流数据的空间特征。CNN 层通过卷积和池化的操作, 进行局部趋势学习, 有效保留交通流的空间特征。神经网络模型层数越多, 网络表现能力则越强。当训练数据不

足时, 更容易发生过拟合。为平衡模型的计算成本, 模型的一维 CNN 层是深度为 2 的完整卷积神经网络。在深度卷积神经网络中, 线性整流层的激活单元用于非线性激活。线性整流函数(Rectified Linear Unit, 简称为 ReLu), 是一种常用的激活函数。不仅可以解决“梯度消失”的问题, 而且可使深层神经网络快速收敛<sup>[14]</sup>。交通流数据具有非线性特征, 所以选择 ReLu 作为整流层激活函数, 提高网络非线性特征的学习能力。

### 1.2 GRU 模型结构

由于交通流的动态性, 交通流的时间特征不断变化, 使得短时交通流表现出更强的相关性<sup>[15]</sup>。由于交通流还存在长期依赖性特征, 如: 上游交通流在拥堵状态下, 需要很长时间, 才能到达下游。因此, 交通流的时间特征建模时, 需要考虑这种长期依赖性。因为 LSTM 可用于模拟交通流的时间特征<sup>[16]</sup>, 通过输入门、遗忘门及输出门 3 个门控机制, 去除或者增加信息, 到 LSTM 单元中, 解决了 RNN 存在的长期依赖, 所以可选择更简单的 RNN-GRU 模型。与 LSTM 相比, GRU 的神经元更少, 而且能实现与 LSTM 相同或者更好的性能<sup>[17]</sup>。

GRU 只有更新门和重置门 2 个门控机制。更新门的作用是决定当前输入信息( $X_t$ )有多少是需要保留的; 重置门的作用是将  $X_t$  和前面传递的信息( $h_{t-1}$ )相结合。这 2 个门控机制使得 GRU 具有交通流的长期依赖特征。

GRU 的计算步骤:

1) 在时间步  $t$ ,  $X_t$  和  $h_{t-1}$ , 经过线性变换, 进入更新门的 Sigmoid 层, 输出激活值  $z_t$ 。重置门与输出门的表达式一样, 但权重矩阵不同, 同样输出一个激活值  $r_t$ 、 $z_t$  和  $r_t$  的值域均为  $[0,1]$ 。其中, 0 代表完全遗忘信息, 1 代表完全保留信息。

$$z_t = \sigma(\mathbf{W}^{(z)}X_t + \mathbf{U}^{(z)}h_{t-1})。 \quad (1)$$

$$r_t = \sigma(\mathbf{W}^{(r)}X_t + \mathbf{U}^{(r)}h_{t-1})。 \quad (2)$$

式中:  $\mathbf{W}$  和  $\mathbf{U}$  为权重矩阵;  $X_t$  为  $t$  时刻的输入向量;  $h_{t-1}$  为时间步  $t-1$  保存的信息。

2) 由 tanh 函数计算当前记忆信息  $\tilde{h}_t$ 。时间步  $t$  的最终记忆信息  $h_t$ , 由  $\tilde{h}_t$ 、 $h_{t-1}$  和更新门激活值  $z_t$  加权得出。

$$\tilde{h}_t = \tanh(\mathbf{W}X_t + r_t\mathbf{U}h_{t-1})。 \quad (3)$$

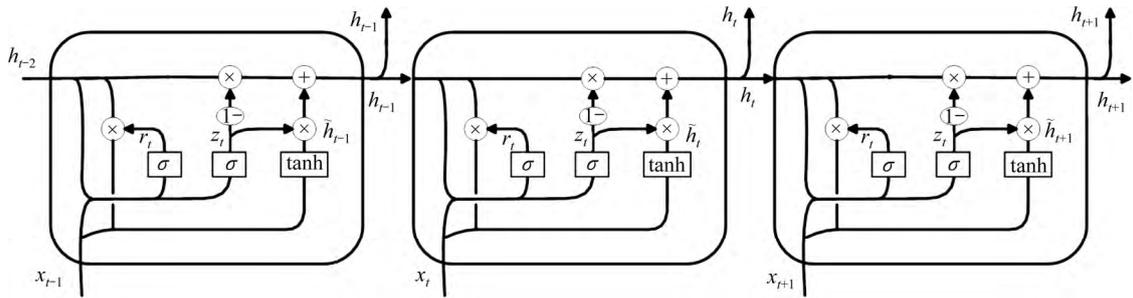


图 1 GRU 结构展开

Fig. 1 GRU expanded structure view

$$h_t = z_t \nabla h_{t-1} + (1 - z_t) \nabla \tilde{h}_t \quad (4)$$

CNN-GRU 模型结构如图 2 所示。

式中： $\nabla$  代表 Hadamard 乘积。

### 1.3 CNN-GRU 模型结构

CNN-GRU 模型利用卷积结构，提取交通流数据中的空间特征，门控循环单元捕获交通流数据的时间特征。在训练阶段时，神经网络从历史数据中学习交通流的时空特征，对交通流的状态和特征进行识别和记忆，得到各个网络层的权重值和偏置量。在预测阶段，将处理好的数据，输入已训练好的模型，得到交通流的预测值。

在考虑计算量和模型深度基础上，选择卷积神经网络层数为 2 层，2 层之间包含一个池化层。模型使用 ReLu 函数，激活 2 个一维卷积层(Conv1D)，提取交通流数据的空间特征。ReLu 函数能保证交通流数据的非负性，并且易于收敛。CNN 每层卷积核数目为 150，卷积核时域窗长度为  $1 \times 10$ 。交通流数据经过卷积层后，使用 2 层 GRU 来提取时间特征，其神经元个数分别为 64 和 32，平衡计算成本和预测精度。Dropout 层设定随机断开 50% 的神经元，可有效缓解模型过拟合现象<sup>[18]</sup>。最终的

## 2 CNN-GRU 模型训练

### 2.1 数据说明及预处理

采用的交通流数据来自 2018 年 9 月-10 月，圣地亚哥区域内标号为 I5-N 高速公路，将其中连续 16 个探测器记录的单向车流量数据，作为样本数据，监测点距离见表 1，空间位置如图 2 所示。数据采集间隔为 30 s，交通量汇总为 5 min。由于采集数据在工作日和周末均表现出较强且不同的周期性模式。因此，本研究只对工作日的工作日所对应的数据作为模型训练集，合计 9 791 个。10 月 22 日-31 日期间的工作日所对应的数据，作为模型测试集，合计 2 303 个。

表 1 监测点间距

Table 1 Distance between monitoring points

编号	距离/km	编号	距离/km
1-2	1.73	9-10	0.84
2-3	0.96	10-11	1.10
3-4	1.18	11-12	1.05
4-5	1.13	12-13	1.12
5-6	1.04	13-14	0.88
6-7	1.13	14-15	0.93
7-8	0.93	15-16	1.12
8-9	1.17	平均	1.09

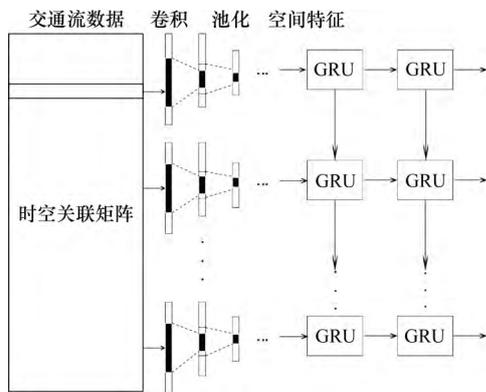


图 2 CNN-GRU 模型结构示意图

Fig. 2 An illustration of structure in CNN-GRU

受外界因素干扰，原始的交通流数据，会出现缺失、异常等状况。本研究利用交通流理论的筛选法，剔除原始数据的异常值。根据交通流在时间上的形似性，对于某一时段缺失的数据，采用前 1 d 或后 1 d 的历史数据替代，将交通流数据归一化到 [0,1] 区间，训练所有模型。

## 2.2 基于时空特征的输入向量

交通流预测实质是一个时间序列的预测, 可以依据过去时段的交通流信息, 预测未来时段的交通流状况。交通流遵循一定的规律, 呈连续流状态, 连续流道路的相关断面相互影响, 使得断面观测到的交通流, 具有时间和空间特征, 需要将交通流量的数据, 处理成带有时间和空间特征的输入矩阵  $\mathbf{A}$ 。假定交通流数据的采样间隔为  $\Delta t$ , 选择的监测点总数为  $p$ , 采集数据的总量为  $n$ , 则所有监测点组成的一组交通流量数据为:

$$\mathbf{A} = (X_1, X_2, X_3, \dots, X_p) \quad (5)$$

式中:  $X_i = \{X_{i,t_1}, X_{i,t_2}, X_{i,t_3}, \dots, X_{i,t_n}\}$ ,  $i=1,2,3,\dots,p$  为第  $i$  个监测点在  $t_n$  时刻的交通流量;  $t_2-t_1=\Delta t$ 。

综合考虑各个监测点在时间与空间维度上的相关性, 设预测模型的交通流量输入为  $\mathbf{X}$ , 则有:

$$\mathbf{X} = \begin{bmatrix} X_{1,t_1} & X_{2,t_1} & \dots & X_{p,t_1} \\ X_{1,t_2} & X_{2,t_2} & \dots & X_{p,t_2} \\ \vdots & \vdots & & \vdots \\ X_{1,t_n} & X_{2,t_n} & \dots & X_{p,t_n} \end{bmatrix}.$$

任意监测点下一时刻的交通流量, 由该监测点和各个监测点的当前时刻及前一个或多个时刻的交通流量数据来预测。训练时, 利用枚举试算的方法, 确定模型预测的时间滞后, 设置为 6, 以实现模型的最低错误率。即: 使用所有监测点前 30 min 的交通流量, 作为定期输入, 预测下一时刻 5 min 的交通流量。

## 2.3 模型训练

利用交叉验证法训练模型, 选择最佳的层数和超参数。训练过程中, 利用损失函数之均方误差 (Mean Squared Error, 简称为 MSE)、Relu 激活函数及 Adam 优化器。Adam 是一种特殊的随机梯度下降 (Stochastic Gradient Descent, 简称为 SGD) 方法, 通过动态调整参数, 充分提取数据的空间特征, 达到快速收敛, 正确学习目标。采用 MSE 作为模型的目标函数, 对比不同超参数的模型误差, 选择最佳方案。其中, 混合模型、CNN、LSTM 及 GRU 模型的 Batchsize 参数为全数据集 288, 能充分学习数据的本质特征, 加快收敛速度, 并且满足预测精度要求。模型交通量的预测时间步长为 3,6,9。3 项评价指标都很小, 步长为 6 时, 平均精度最高。自

回归移动平均模型 (Autoregressive Integrated Moving Average Model, 简称为 ARIMA) 超参数选择 (1,1,1)。

在 Python 语言开发环境下, 所有神经网络都基于 Keras 框架, 进行搭建和训练。基于 Keras, 可将训练模型保存到本地, 在预测阶段可直接调用进行预测实验。根据预测精度自适应调整模型参数, 具备可移植性和灵活性。

## 3 CNN-GRU 模型预测结果和分析

### 3.1 单个监测点模型预测效果

基于训练好的混合模型和对照模型, 随机选取 10 月 31 日编号为 4 的监测点数据作为预测样本, 模型预测结果如图 3 所示。

从图 3 可以看出, CNN-GRU 模型预测结果最好, 能更好的预测出实际交通流的变化趋势, 预测值和实际值的误差较小, 平均误差低至 1 veh/5 min。其中, 交通流在凌晨 (0-50 时段) 比较平稳时, 平均误差为 0。但在早晚高峰 (50-220 时段) 时, 交通流变化剧烈, 预测误差会相应出现波动, 特别是交通流峰值处, 最大误差达到了 5 veh/5 min。从图 4 中还可以看出, CNN 模型对于交通流的峰值预测效果偏低, GRU 和 LSTM 模型对于交通流变化趋势的预测效果较差。混合模型中, 利用 CNN 和 GRU 的优点, 识别了短时交通流数据的空间局部特征, 并且保留了挖掘时间特征的能力, 该模型有效提高了预测精度。

为了更好地分析预测结果, 选取均方根误差  $e_{RMSE}$ 、平均绝对误差  $e_{MAE}$  和平均绝对百分比误差  $e_{MAPE}$  作为评价指标衡量模型的性能, 其中, MAPE 是比较不同预测模型准确性的常见评价指标之一。RMSE、MAE 和 MAPE 三者定义如下:

$$e_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (6)$$

$$e_{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (7)$$

$$e_{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (8)$$

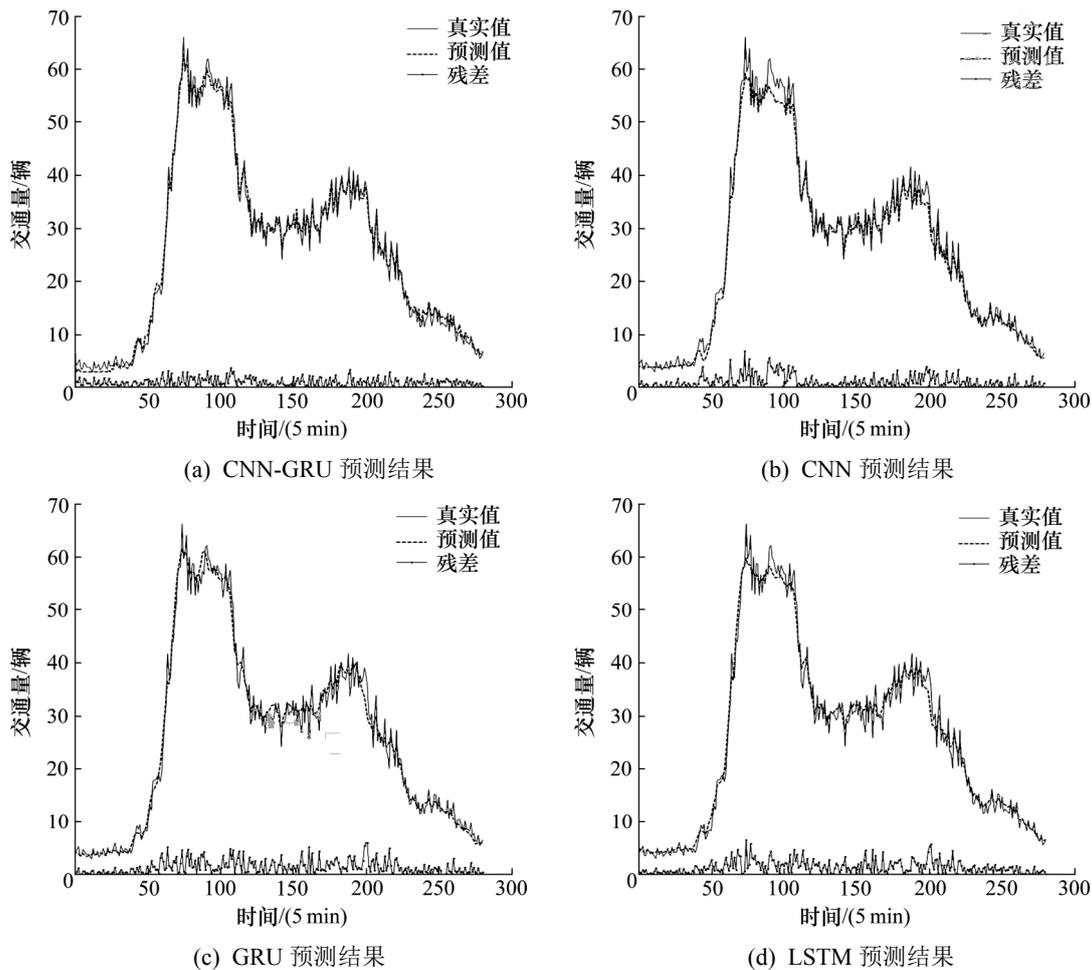


图3 模型的预测结果和误差对比

Fig. 3 Prediction results and error comparison of models

式中： $y_i$  为实际交通流量； $\hat{y}_i$  为预测交通流量； $n$  为样本数量。

4号监测点的预测结果评价见表2。CNN-GRU模型与其他模型相比，在3个评价指标中，均为最低值，预测精度最高。通过神经网络各模型的评价指标对比，CNN-GRU、GRU和LSTM3种神经网络模型评价指标，基本处于同一水准。而ARIMA

表2 四种模型在4号监测点对比RMSE、MAE和MAPE评价指标

Table 2 The comparison of RMSE, MAE and MAPE of four models at No. 4 monitoring point

模型	评价指标		
	RMSE	MAE	MAPE/%
CNN-GRU	1.32	1.08	7.31
GRU	2.23	1.46	8.66
LSTM	2.15	1.61	9.11
CNN	5.39	4.01	10.39
ARIMA	6.95	5.59	11.13

模型的MAPE，高达11.13%，相比其他神经网络模型性能最差。可见神经网络的多隐含层结构，能有效提取交通流数据的特征。此外，混合模型的3项指标最低，MAPE相比LSTM模型减小了19.7%，预测精度大幅提升。

### 3.2 全路段模型预测效果

为了验证CNN-GRU预测模型的适用性，选取10月31日16个监测点的数据，作为预测样本，模型预测结果如图4所示。从图4可以看出，路段交通流随时间变化的情况。图4(b)预测交通量与图4(a)实际交通量的变化趋势一致，所提出的模型在整条实验路段上适用性较好。

为了进一步验证CNN-GRU模型的预测精度，本试验构建3个对比模型：CNN、GRU和LSTM模型。2018年9月-10月16个采样点的预测评价结果，见表3。从表3可以看出，CNN-GRU模型与GRU、LSTM及CNN3种神经网络模型的 $e_{MAPE}$ 相比，CNN-GRU模型的精度指标最高，均优于其他

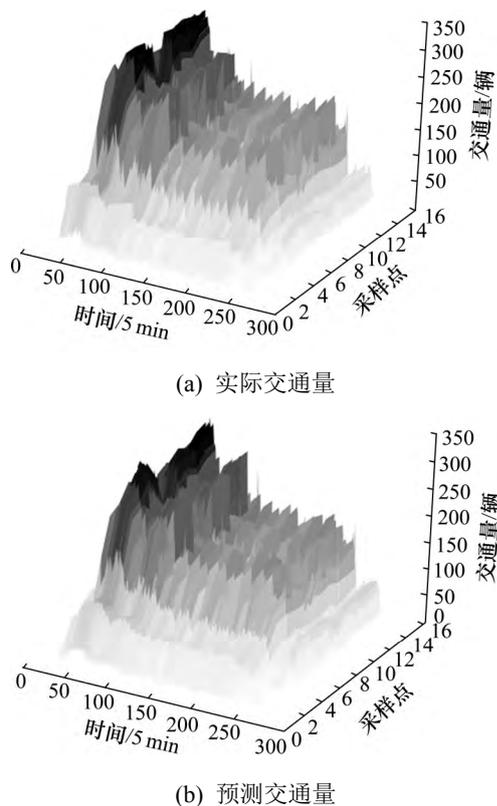


图 4 CNN-GRU 模型全天的预测结果

Fig. 4 Forecast results of all day in CNN-GRU model

预测模型。其中, 相比 GRU 模型, MAPE 减少了 12%; 相比 LSTM 模型, MAPE 减少了 13%; 相比 CNN 模型, MAPE 减少了 23%。表明: 本研究所提出的 CNN-GRU 模型, 预测精度最好。CNN 的预测精度比 GRU 更好, 也进一步验证了 CNN 提取短时交通流空间特征的强大能力。如果与其它体系结构相结合使用, 其实用性更强。

表 3 全路段预测结果评价

Table 3 Evaluation of all road segment prediction results

模型	评价指标		
	RMSE	MAE	MAPE/%
CNN-GRU	2.63	2.11	9.05
GRU	3.53	2.71	10.26
LSTM	3.67	2.79	10.46
CNN	7.09	5.62	11.80

## 4 结语

基于深度神经网络理论, 提出一种卷积神经网络和门控循环单元神经网络相结合的 CNN-GRN 预测模型, 对该模型进行训练与预测, 得到的结论为:

1) 利用 Python 语言, 完成 CNN-GRU 模型的

逐层构建和参数调优, 实现基于 CNN-GRU 模型的短时交通量的精确预测。

2) 利用 PeMS 的交通量数据训练模型, 并进行对比测试单个监测点和全路段的预测结果表明: CNN-GRU 模型能够对交通流的状态和特征进行识别和记忆, 充分提取交通流数据的时间和空间特征, 提高模型的预测精度。

3) CNN-GRN 与 CNN、GRN 及 LSTM 3 种模型相比, 3 项评价指标均为最低, 预测精度最高, 具备较好的适用性。

4) 实例数据证明, CNN-GRU 模型能更精确预测交通流的短时变化, 可为交通管理与控制提供有效依据。

本实验仅仅是针对某个路段, 未来应该扩展到城市路网中, 如此更符合城市交通智能化需要。

## 参考文献(References):

- [1] 张培林, 吴宏宇, 莫杨辉, 等. 大城市综合交通适应性评价[J]. 交通科学与工程, 2018, 34(4): 80-85. (ZHANG Pei-lin, WU Hong-yu, MO Yang-hui, et al. Adaptability evaluation of the comprehensive traffic in large cities, 2018, 34(4): 80-85. (in Chinese))
- [2] Clark S. Traffic prediction using multivariate nonparametric regression[J]. Journal of Transportation Engineering, 2003, 129(2): 161-168.
- [3] Oh S, Byon Y J, Jang K, et al. Short-term traveltime prediction on highway: A review on model-based approach[J]. KSCE Journal of Civil Engineering, 2017, 35(1): 4-32.
- [4] 刘兆惠, 李倩, 王超, 等. 基于小波卡尔曼滤波的高速公路交通数据融合去噪算法研究[J]. 公路工程, 2018, 43(6): 91-96. (LIU Zhao-hui, LI Qian, WANG Chao, et al. Research on denoising algorithm of expressway traffic data fusion based on wavelet Kalman filter[J]. Highway engineering, 2018, 43(6): 91-96. (in Chinese))
- [5] Hamed M M, Al-Masaeid H R, Said Z M B. Short-term prediction of traffic volume in urban arterials[J]. Journal of Transportation Engineering, 1995, 121(3): 249-254.
- [6] 张涛, 陈先, 谢美萍, 等. 基于 K 近邻非参数回归的短时交通流预测方法[J]. 系统工程理论与实践, 2010, 30(2): 76-384. (ZHANG Tao, CHEN Xian, XIE Mei-ping, et al. K-NN based nonparametric regression method for

- short-term traffic flow forecasting[J]. *Systems Engineering-Theory & Practice*,2010,30(2):76-384.(in Chinese))
- [7] Hu W,Yan L,Liu K, et al. Pso-Svr: A hybrid short-term traffic flow forecasting method[C]// *IEEE International Conference on Parallel & Distributed Systems*. IEEE, 2016.
- [8] Du S,Li T,Gong X,et al.Traffic flow forecasting based on hybrid deep learning framework[C]// *2017 12th International Conference on Intelligent Systems and Knowledge Engineering*. IEEE, 2018.
- [9] Lv Y, Duan Y, Kang W, et al. Traffic flow prediction with big data: a deep learning approach[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(2): 865-873.
- [10] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*,1997,9(8):1735-1780.
- [11] Polson N G, Sokolov V O. Deep learning for short-term traffic flow prediction[J]. *Transportation Research Part C Emerging Technologies*,2017,79:1-17.
- [12] Yu H, Wu Z, Wang S, et al. Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks[J].*Sensors*,2017,17 (7):1501.
- [13] 刘明宇,吴建平,王钰博,等.基于深度学习的交通流量预测[J]. *系统仿真学报*, 2018, 30(11): 77-82, 91. (LIU Ming-yu,WU Jian-ping,WANG Yu-bo,et al.Traffic flow prediction based on deep learning[J].*Journal of System Simulation*,2018,30(11):77-82,91.(in Chinese))
- [14] Zeiler M D,Ranzato M,Monga R,et al.On rectified linear units for speech processing[C]// *IEEE International Conference on Acoustics*. IEEE, 2013.
- [15] Wu Y, Tan H, Qin L, et al. A hybrid deep learning based traffic flow prediction method and its understanding[J]. *Transportation Research*,2018,90(MAY):166-180.
- [16] Duan Y, Lv Y ,Wang F. Travel time prediction with LSTM neural network[C]// *IEEE International Conference on Intelligent Transportation Systems*. IEEE, 2016.
- [17] Cho K, Van Merriënboer B, Gulceher C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. *ArXiv Preprint*, 2014, 1406:1078.
- [18] Sricastava N, Hinton G, Krizhevsky A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. *Journal of Machine Learning Research*, 2014, 15(1): 1929-1958.