

# 基于深度学习的短时交通量预测研究综述

代亮<sup>1</sup> 梅洋<sup>1</sup> 钱超<sup>1</sup> 孟芸<sup>1</sup> 吕金明<sup>2</sup>

(长安大学电子与控制工程学院 西安 710064)<sup>1</sup> (IBM 中国系统与科技开发中心 西安 710068)<sup>2</sup>

**摘要** 短时交通量预测是智能交通领域的研究热点,对交通控制与管理具有重要的意义。传统的交通量预测方法难以准确地描述交通量数据内部的本质特征,而深度学习通过其深层结构,能够学习到交通量数据内部复杂的多因素耦合结构,进而对交通量做出更精准的预测,这也使得深度学习成为当前短时交通量预测领域的研究热点。文中首先介绍了传统交通量预测方法和深度学习的研究现状;然后按照生成型和判别型深度结构对现有基于深度学习的短时交通量预测方法进行分类,并总结了深度学习在短时交通量预测研究领域的主要方法,对其性能进行了对比研究;最后对深度学习在短时交通量预测领域存在的问题和发展趋势进行了探讨。

**关键词** 短时交通量预测,交通控制与管理,深度学习,生成型深度结构,判别型深度结构

**中图分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.03.005

## Survey on Short-term Traffic Flow Forecasting Based on Deep Learning

DAI Liang<sup>1</sup> MEI Yang<sup>1</sup> QIAO Chao<sup>1</sup> MENG Yun<sup>1</sup> LV Jin-ming<sup>2</sup>

(School of Electronics and Control Engineering, Chang'an University, Xi'an 710064, China)<sup>1</sup>

(IBM China Systems and Technology Laboratories, Xi'an 710068, China)<sup>2</sup>

**Abstract** Short-term traffic flow forecasting is a hot topic in the field of intelligent transportation, which is of great significance in traffic control and management. The traditional traffic flow forecasting methods are difficult to describe the internal characteristics of the traffic data accurately. Deep learning can learn the internal complex multivariate coupled structure of the traffic flow data through its deep structure and then make a more accurate forecasting of the traffic flow, which makes deep learning a hot topic in the current traffic flow forecasting field. Firstly, the traditional traffic flow forecasting methods and the current research status of deep learning were briefly introduced. Then the methods of short-term traffic flow forecasting based on deep learning were classified according to generative deep architecture and discriminative deep architecture. This paper also summarized the main methods of deep learning in the field of traffic flow forecasting and compared their performance. Finally, the existing problems and development directions of deep learning in short-term traffic flow forecasting were discussed.

**Keywords** Short-term traffic flow forecasting, Traffic control and management, Deep learning, Generative deep architecture, Discriminative deep architecture

## 1 引言

随着社会经济的发展和城市规模的扩大,道路里程数与汽车保有量飞速增长,交通拥堵日益成为制约城市与经济发展的主要问题之一。构建智能交通系统(ITS)能够有效缓解道路拥堵,缩短出行时间,降低环境污染,提高交通安全性。作为 ITS 的重要研究内容之一,交通量预测能够作为交通决策的重要依据,可有效地应用于交通规划、交通控制和交通诱导。实现实时有效的交通控制需要对路网中的交通量进行准确、实时的预测及研判,及时发现路网中存在的交通问题,以提高道路的运行效率。

交通量预测按照预测间隔时长可分为中长期和短时交通量预测。中长期交通量预测的预测间隔以天、月甚至年为单单位,通常用于交通规划等需要对交通量进行宏观预测的情形。短时交通量预测一般指预测间隔为 5~15 min 的预测。由于预测间隔越小,交通量受随机因素的影响越大,导致其不确定性增强,使得交通量表现出不稳定性,增加了预测的难度。由于短时交通量的随机性强,预测难度大,在如今愈发复杂的交通状况下,短时交通量的准确、实时预测仍然是一个亟待解决的难题。

深度学习作为一个较新的研究领域,由 Hinton 在 2006 年首次提出<sup>[1]</sup>。自此之后,深度学习成为了人工智能领域的

到稿日期:2018-04-27 返修日期:2018-08-24 本文受国家自然科学基金(61701044),中国博士后特别资助项目(2015T80998),陕西省自然科学基金(2016JQ6067),陕西省博士后科研项目(2014-074),IBM 公司合作项目(SOWCHD201610)资助。

代亮(1981-),男,博士,副教授,CCF 会员,主要研究方向为交通数据分析,E-mail:ldai@chd.edu.cn(通信作者);梅洋(1992-),男,硕士生,主要研究方向为人工智能与深度学习;钱超(1984-),男,博士,讲师,主要研究方向为交通数据挖掘与分析;孟芸(1987-),女,博士,讲师,主要研究方向为智能交通系统;吕金明(1982-),男,高级工程师,主要研究方向为大数据技术。

一个热门研究方向,受到学术界与工业界的广泛关注。深度学习研究将多个数据层特征进行组合,进而形成多维特征的应用层语义抽象,从而减少了传统机器学习方法需要主观先验经验来设计特征的人为因素影响,可使发现数据分布式特征表示的过程智能化。深度学习在计算机视觉、机器翻译、语音识别等众多领域都取得了突破性的进展。为了解决当前交通状况下短时交通量的准确预测问题,2014年深度学习首次被应用于短时交通量预测中。本文对近年来深度学习在短时交通量预测领域中的研究进行了系统性总结,并且对已有方法及其性能进行对比与分析,同时对当前该领域存在的问题和未来的发展方向进行了探讨。

## 2 传统短时交通量预测方法

短时交通量预测是交通控制与诱导的基础,是智能交通系统的重要研究内容之一。随着交通量预测相关研究的不断推进,研究人员已经提出了大量短时交通量预测方法。根据模型结构,这些短时交通量预测模型可分为参数模型和非参数模型。

### 2.1 参数模型

时间序列模型在参数模型中有着广泛的应用。在交通量预测研究领域,常用的时间序列模型有自回归模型(AR)、移动平滑模型(MA)、自回归滑动平均模型(ARMA)、自回归求和滑动平均模型(ARIMA)。其中应用最广泛的模型是1979年提出的ARIMA模型,之后Levin等发现了ARIMA(0,1,1)模型用于交通量预测时的统计显著性最佳<sup>[2-3]</sup>。另外,ARIMA模型具有多种变体,如Williams等<sup>[4]</sup>考虑到交通量周期性的特点,提出了季节性ARIMA模型来提高预测准确率;Lee等<sup>[5]</sup>通过对比不同时间序列模型的短时交通量预测结果,发现子集ARIMA的预测结果具有更高的准确率和稳定性。另一种常用的参数模型是卡尔曼滤波模型<sup>[6]</sup>,该模型对预测因子的选择灵活,可处理平稳或非平稳数据。研究发现,基于卡尔曼滤波的短时交通量预测模型在实验中取得了良好的预测结果<sup>[7]</sup>。

参数模型的优点在于模型简单,可解释性强,运算速度快;其缺点也同样突出,即交通流非线性、随机性强的特征,使得基于参数模型的交通量预测方法的准确率较低、误差较大。

### 2.2 非参数模型

针对参数模型无法描述交通流非线性和随机性强的特征的问题,非参数模型成为了一条有效的解决途径,如支持向量机模型<sup>[8]</sup>、非参数回归模型<sup>[9-10]</sup>、支持向量回归模型<sup>[11]</sup>、高斯过程模型<sup>[12]</sup>、K最近邻模型<sup>[13-14]</sup>、贝叶斯组合模型<sup>[15]</sup>等。人工神经网络由于具有捕捉复杂非线性特征的能力,因此适用于对交通量进行预测。Smith和Demetsky利用反向传播(BP)神经网络对短时交通量进行预测,从均方根误差和平均绝对误差这两个评价指标来看,BP神经网络较ARIMA的预测性能有较大的提升<sup>[16]</sup>。Messer等利用径向基函数神经网络(RBFNN)预测短时交通量,研究表明RBFNN较BP神经网络具有更好的泛化能力和更快的学习速度<sup>[17]</sup>。

非参数模型较为复杂,算法收敛较慢,其学习算法采用了经验风险最小化原理,无法保证期望风险最小化,这使得模型容易产生过拟合问题。另外,模型的优化较困难,并且容易陷入局部最优解。

## 3 基于深度学习的短时交通量预测方法

交通流具有随机性和非线性特征,传统短时交通量预测方法由于其浅层的结构,无法对特定函数族进行有效的表征以获取数据集的本质特征<sup>[18]</sup>。另外,对于高维度数据集,运用传统预测方法会造成维度灾难问题<sup>[19]</sup>,因此传统预测方法倾向于忽略额外变量,这样就无法充分利用交通数据的多属性特点,继而影响对现实交通量预测问题的建模。而深度学习能够弥补传统短时交通量预测方法的不足。

深度学习是机器学习领域中一个较新的研究方向,它是具有多层表征的表征学习方法。在每一层中,通过简单但非线性的模块将原始数据转换为高层更为抽象的表征,在经过足够多的转换后,即可实现复杂且高度灵活的函数逼近<sup>[20]</sup>。单层学习算法可以提取出显著的特征,但由于其学习能力的限制,第一层提取的特征被视为低级特征,第二层基于同样的原理,将第一层的输出特征作为输入来提取更高级的特征,利用这种办法,最终将会得到更高层的特征,从而更好地表示数据内部的复杂结构<sup>[21]</sup>。

2006年,Hinton等<sup>[1]</sup>提出了基于深度信念网络(DBN)的快速学习算法。该算法使用非监督逐层贪婪预训练方法来获得模型的权重参数,通过多层表征学习得到更好的数据表征描述,同时逐层训练方法的应用也降低了深度神经网络的训练难度。这一研究成果改变了原先深层网络无法被有效训练的局面,为使深度学习技术的飞速提升奠定了基础。随着可用交通量数据的不断增多和计算机软硬件性能的提高,深度学习成为了短时交通量预测领域中最活跃的研究方向。根据深度学习不同的结构、训练方法等因素,可将其分为3类<sup>[22]</sup>。

1)生成型深度结构。生成型深度结构可描述数据的高阶相关特征,形成观测数据及其标签的联合概率分布模型,能够对先验概率和后验概率进行估计。通过贝叶斯准则,由生成型模型可得到判别型模型。

2)判别型深度结构。判别型深度结构可提供对模式分类的判别能力,它描述了观测数据的条件概率分布,对后验概率进行估计。判别型深度结构可直接判别特征的类别。

3)混合型深度结构。混合型深度结构的学习过程包括生成型和判别型两部分,由生成单元与判别单元组合而成。利用生成单元对模型强大的表达能力和判别单元高效的判别能力<sup>[23]</sup>,混合型深度结构具有强大的模型判别能力。

本文基于上述对深度学习的结构分类,对现有基于深度学习的短时交通量预测方法进行分类分析,并研究其特点及优势。其中,生成型模型的学习收敛速率较判别型模型更快,当样本容量较大时,生成型模型能够更快地收敛到最优解,提高了训练效率。判别型模型直接学习观测数据的条件概率分布,由于其直接面对预测问题,因此模型的预测准确率往往较生成型模型高。目前暂未有利用混合型深度结构进行短时交通量预测的相关研究。

### 3.1 生成型深度结构

生成型深度结构能够描述数据的高阶相关特性。基于生成型深度结构的短时交通量预测模型包括深度信念网络、自编码器及其扩展模型。

3.1.1 深度信念网络

深度信念网络 (DBN) 由多个受限玻尔兹曼机 (RBM) 堆叠组成,每一层的输出作为下一层的输入。RBM 是包含一层可视层和一层隐藏层的无向概率图模型,其可视层和隐藏层之间的节点全连接,层内节点不允许连接,其模型拓扑结构为一个完全偶图,如图 1 所示。RBM 是一种基于能量的模型,其目的是使模型能够最大概率地得到符合条件的样本<sup>[24]</sup>。Roux 等从理论上证明了,只要隐藏层单元足够多,RBM 就能够拟合任意的离散分布<sup>[25]</sup>。RBM 的特点使其成为了 DBN 的组成单元。

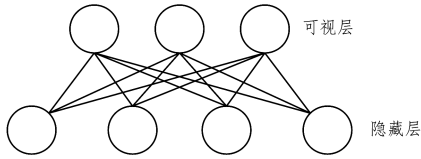


图 1 RBM 结构  
Fig. 1 Structure of RBM

在引入 DBN 之前,深度模型被认为学习过程较慢,难以优化。Hinton 等提出了一种非监督逐层贪婪预训练方法来训练 DBN<sup>[1]</sup>,这种无监督的预训练方法通过更好地初始化权重参数,有效地解决了深度神经网络学习过程慢、易陷入局部最优解的问题;进而采用有监督的训练方法对模型参数进行微调,使模型具有最优化的权重参数。

DBN 算法效率的提高,使其成功应用于短时交通量预测中。Huang 等首次将深度学习引入到短时交通量预测的研究中,模型结构如图 2 所示,该深度结构由 DBN 及其顶部添加的一个多任务回归层两部分组成<sup>[26]</sup>。DBN 可以无需先验知识对数据特征进行学习,可以用来作为特征学习模型,将多任务回归层作为预测模型,利用监督学习对交通量进行预测。该深度结构整合了多任务学习来提升模型的预测性能,利用同一个模型对路段和站点的交通量进行预测。由于路段和站点的交通量具有高度相关性,通过共享权重参数可以提高统计强度,从而改善模型的泛化能力和泛化误差范围<sup>[27]</sup>。实验表明,该模型较当时的最高预测准确率提高了 5%。

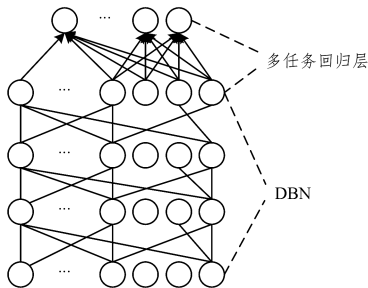


图 2 整合了多任务学习的深度信念网络  
Fig. 2 DBN integrated with multitask learning

文献<sup>[28]</sup>考虑了交通量、天气因素和从社交媒体中提取的事件数据对交通量预测的影响,利用 DBN 分别对两种类型的数据进行交通量预测,再利用 Dempster-Shafer 理论 (DEST)<sup>[29]</sup>将二者的预测结果融合。由于基于事件的数据不是连续产生的,并且具有不可靠性,因此该模型将历史交通量与天气数据这种连续数据流作为主要信息源,而基于事件的数据只有在支持预测结果时才被作为交通量预测的影响因

素。将多种数据源信息相融合,能够提高用于交通量预测的数据表达的充分性,降低噪声数据对模型的影响,有效提高交通量预测的准确率。

文献<sup>[30]</sup>同样利用了数据融合方法来提高交通量预测的准确率和稳定性,其预测流程如图 3 所示。在该文献中,预测模型 1 研究了交通量的自相关性,以此来确定获得高预测准确率所需的延迟间隔。由于交通网络是典型的非线性系统,因此其采用了斯皮尔曼等级相关系数来计算交通量的自相关系数。预测模型 2 研究了交通量与天气情况的互相关性,用于选择对交通量影响大的天气变量。进而,将两个 DBN 预测模型的预测结果进行数据融合。该模型考虑到了交通数据的多属性特点,实验结果表明利用决策层融合可以有效提高预测准确率。

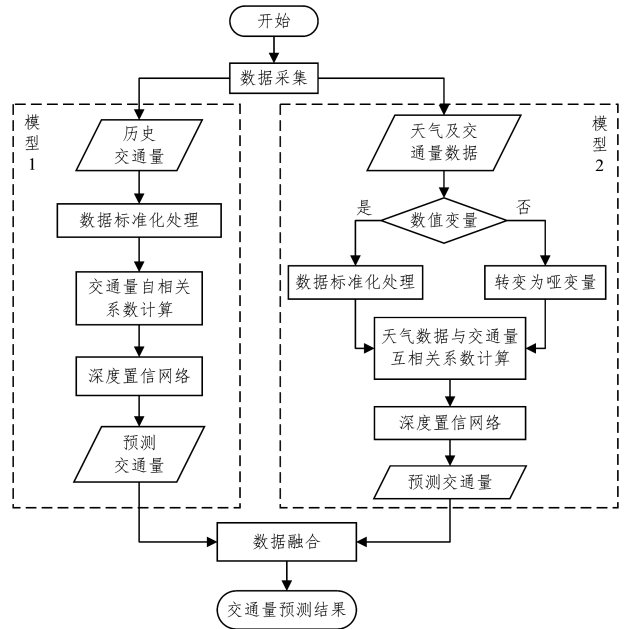


图 3 对多源数据进行融合的交通量预测流程

Fig. 3 Traffic flow forecasting process using multi-source data fusion

文献<sup>[31]</sup>采用了 DBN-SVR 模型,即在多个 RBM 模型对数据进行特征提取后,再在顶部连接一个支持向量回归分类器 (SVR) 来预测交通量。SVR 适合于处理非线性数据,但对具有趋势项的时间序列数据的处理能力较差。因此,为了提高预测的准确率,文献<sup>[31]</sup>利用差分原理对交通量数据进行预处理,以消除趋势项。实验结果表明,相较于原始数据,去趋势的数据将预测准确率提高了 22.34%。从实验结果可以看出,交通数据的趋势项对预测结果有较大的影响。

文献<sup>[32]</sup>通过对比 G-B DBN 和 B-B DBN 模型,发现前者在交通量预测方面的准确率更高。这两种 DBN 模型的差异在于其组成单元 RBM 不同,其中一种 RBM 的可视单元为高斯单元,隐藏单元为二值单元 (G-B RBM)<sup>[33-34]</sup>,由 G-B RBM 构成的 DBN 模型称作 G-B DBN,其结构如图 4 所示;另一种 RBM 的可视单元、隐藏单元均为二值单元 (B-B RBM),由 B-B RBM 构成的 DBN 模型称作 B-B DBN。实验证明,G-B DBN 的预测准确率较 B-B DBN 高 0.85%,其原因是 B-B RBM 的隐藏单元和可视单元的状态均为 0 或 1,这种方式会引起采样噪声,而 G-B RBM 能够有效减少采样噪声,提高预测准确率。另外,B-B RBM 可视层的输入均是二值的,不适

用于模拟现实中交通量这种连续值的问题。G-B RBM 加入了高斯噪声的连续值,这样可以更有效地对现实交通量预测问题进行建模。

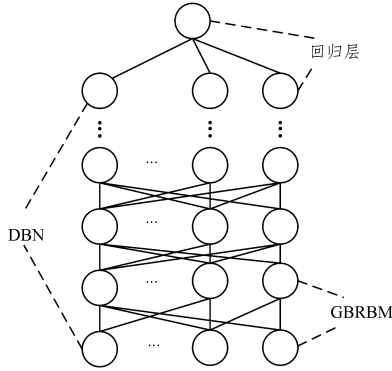


图4 底层由 G-B RBM 组成的 DBN 的架构  
Fig. 4 Structure of DBN consisting of G-B RBM

DBN 是最早应用深度结构训练的模型之一,同时也是最早应用于交通量预测的深度模型。现有研究在该模型的基础上利用多任务学习、数据融合等技术减小模型泛化性能不佳的风险,以提升模型的预测性能。

### 3.1.2 自编码器及其扩展模型

自编码器可被看作由一个编码器和一个解码器两部分组成,常用于数据降维和特征提取,其结构如图 5 所示。函数  $h=f(x)$  表示的编码器将输入  $x$  映射到隐藏层  $h$ ,隐藏层  $h$  通过编码器学习到输入信息的特征,得到输入信息的编码表示。函数  $z=g(h)$  表示的解码器将隐藏层  $h$  映射到输出  $z$ 。自编码器尝试学习一个恒等函数  $x=g(f(x))$ ,使得输出  $\hat{x}$  等于输入  $x$ 。当增加约束条件,使得自编码器的隐藏层  $h$  的维度小于输入层  $x$  时,便可强制自编码器学习数据的压缩表示,提取数据最显著的特征<sup>[35]</sup>。当隐藏层  $h$  的维度大于输入层  $x$  时,可以给隐藏层神经元增加稀疏性约束条件,同样可以提取有用的特征<sup>[36]</sup>。

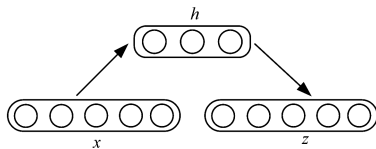


图5 自编码器的结构  
Fig. 5 Structure of autoencoder

堆栈自编码器(SAE)由若干自编码器堆叠而成,使模型能够进行多层表征学习,达到提取数据显著特征的能力,从而挖掘交通量数据中的时空相关性,其结构如图 6 所示。

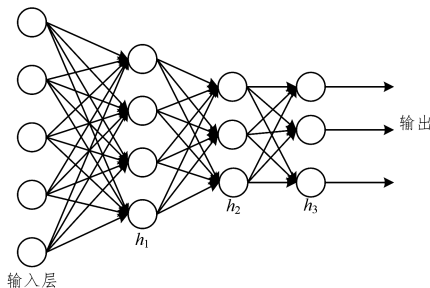


图6 SAE 的结构  
Fig. 6 Structure of SAE

SAE 的训练方法也可以采用非监督逐层贪婪预训练方法,即利用输入层的原始输入训练第一个自编码器,得到原始输入的一阶特征表示,然后将隐藏层  $h_1$  作为下一个自编码器的输入来学习二阶特征表示,以此来学习到更高阶的特征表示。文献[37-40]分别采用了不同深度结构、不同训练方法的 SAE,提出了具有不同特点的交通量预测算法。

Lv 等首次使用以自编码器为建构模块的深度结构进行交通量预测<sup>[37]</sup>。SAE 模型在 4 组实验中的预测准确率均超过了 93%,相比于其他具有浅层结构的算法,在预测准确率、绝对误差以及稳定性方面具有优势。实验结果表明,SAE 能够有效地发现潜在的交通量特征表示。

文献[38]采用 SAE 深度模型,更进一步地评估了深度学习模型在不同时间段的交通量预测中的表现。鉴于工作日与非工作日的交通量数据模式不同,研究人员分别对工作日与非工作日采用了不同超参数的 SAE 模型,对白天(6 a. m. - 8 p. m.)与夜间(8 p. m. - 6 a. m.)两个时间段进行了交通量预测实验。实验表明,相对于非工作日,无论是白天还是夜间,工作日的预测模型为了达到最佳效果,趋向于选取较多的隐藏层数,说明当数据模式越复杂时 SAE 模型所需要的隐藏层数越多。

Yang 等<sup>[39]</sup>提出了基于 Levenberg-Marquardt 算法的堆栈自编码器模型(SAE-LM)。该模型利用指数平滑<sup>[41]</sup>和 Levenberg-Marquardt(LM)<sup>[42]</sup>算法来对神经网络进行训练,其中指数平滑方法用来消除随机因素对交通量数据造成的影响;而 LM 算法在训练神经网络时,对非线性方程的求解更为稳定且具有良好的收敛性,从而更有效地使函数在函数参数空间中最小化。另外,在参数选取的实验中,研究人员采用了 Taguchi 方法<sup>[43]</sup>来获取交通量预测模型的优化结构,使得实验次数相较于尝试法大大减少,从而提高了交通量预测模型设计的高效性。SAE-LM 是解决波动数据预测问题最合适的方法,预测准确率达到 90% 左右。但当交通数据为高度平滑分布时,该模型的预测准确率较低。

Zhou 等<sup>[40]</sup>构建了多种不同深度结构、不同初始值的 SAE 模型,再利用一个基于  $\delta$ -agree AdaBoost 回归的集成方案来提升学习模型。由于交通量这种低维度数据存在大量的鞍点<sup>[44]</sup>,单个用于预测的 SAE 网络容易陷入局部最优解。利用 AdaBoost 算法<sup>[45]</sup>可以减小选择最差预测结果的风险,能够解决选取最优模型参数和结构的不确定性。同时,AdaBoost 算法可有效防止模型过拟合。重构误差反映了模型重构输入值的能力,而预测误差反映了模型的泛化能力,因此研究人员利用预测误差,通过重排训练数据来重新训练 SAE,从而达到提高交通量预测准确率的目的。

自编码器的编码器和解码器都是前向神经网络,这两部分都能从深度结构中获取到更为强大的表示能力。通过自编码器的堆叠,SAE 能够对交通量数据进行多层表征学习,以学习到序列数据间的潜在特征。现有研究基于 SAE 的特点改进网络的训练算法,引入集成学习来提高模型的鲁棒性,降低预测偏差。

### 3.2 判别型深度结构

判别型深度结构能够描述数据的条件概率分布。基于判

别型深度结构的短时交通量预测模型包括循环神经网络及其扩展模型、卷积神经网络。

### 3.2.1 循环神经网络及其扩展模型

循环神经网络(RNN)<sup>[46]</sup>由具有非线性动态的高维度隐藏状态组成,能够记忆并处理过去的信息<sup>[47]</sup>,并从序列化和时序数据中学习特征和长期依赖关系。自 Elman<sup>[48]</sup>展示了 RNN 具有学习复杂时空内部特征的能力以来,RNN 便成为了解决交通序列数据问题的最佳选择<sup>[49]</sup>。前向神经网络(FFNN)中,输出值仅由本时刻的输入决定,而 RNN 通过循环连接隐藏层节点来捕捉序列化数据中的动态信息,这种循环连接可将之前几个隐藏层的数据作为当前的输入,从而使网络的内部状态可以保留之前输入数据的信息。

RNN 通过循环连接获得了处理具有顺序依赖性数据的能力,但当处理长期及远距离依赖问题时,网络的训练将会出现梯度消失或梯度爆炸的问题<sup>[50-51]</sup>,而长短期记忆(LSTM)模型可以有效解决这一问题<sup>[52]</sup>。如图 7 所示,LSTM 利用记忆单元代替普通 RNN 的隐藏单元。其中,记忆单元由输入门  $i_t$ 、遗忘门  $f_t$  和输出门  $o_t$  这 3 个门控单元组成。这些单元控制了隐藏神经元间的信息流,其中输入门控制当前输入信息存储至记忆单元的比例,遗忘门控制之前时间步中单元状态信息留存的比例,输出门可以消除权重冲突<sup>[53]</sup>。这些特性使得 LSTM 模型相较于 RNN 可以更长期地保存信息和传播误差,能更有效地处理长期及远距离依赖关系。文献<sup>[54-60]</sup>利用 LSTM 模型对短时交通量预测进行了研究与实验。

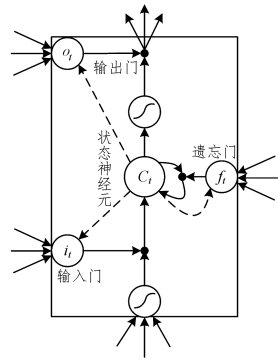


图 7 LSTM 结构图

Fig. 7 Structure of LSTM

为了增强路网间时空关联特性的描述表达,Zhao 等<sup>[54]</sup>提出了起终点关联(ODC)矩阵,即利用观测点  $i$  在时间  $t$  的交通量  $x_{i,t}$ ,计算出不同交通观测点交通量的相关系数构成的 ODC 矩阵,以此体现出路网中不同路段连接的时空相关性。ODC 矩阵作为参数,通过全连接层和向量生成,为 LSTM 网络中的记忆单元生成一个新的时间序列  $X_{i,t}$ ,它能够帮助 LSTM 网络捕捉到交通量变化的特征,其二维结构如图 8 所示。实验分别对代表交通量高、中、低的 3 个观测点进行了时间间隔为 15min 的预测,平均相对误差分别为 6.41%, 6.05%, 6.21%。从数据中可以看出,这种整合了 ODC 矩阵的 LSTM 网络模型的预测准确率较高,且模型泛化能力较强,具有较好的鲁棒性。

时间间隔

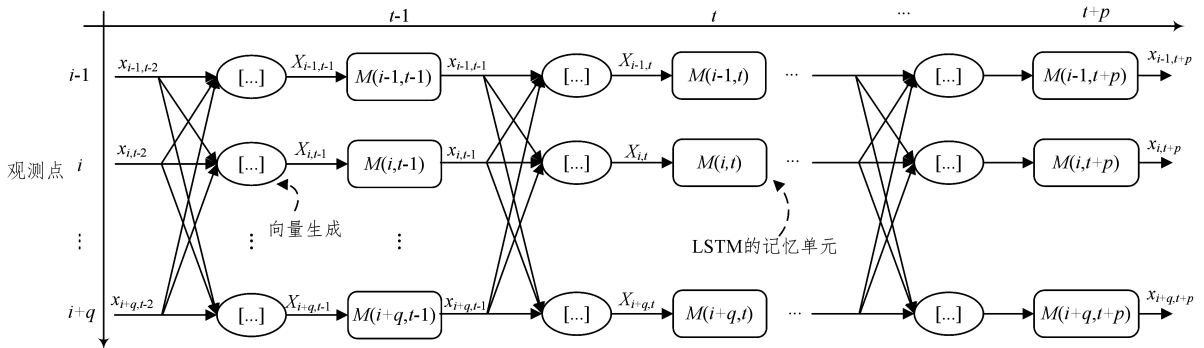


图 8 二维 LSTM 网络结构

Fig. 8 Two-dimensional LSTM network structure

现有的利用深度学习进行交通量预测的文献中,绝大多数都忽略了对交通量预测有很大影响的非交通数据源<sup>[61-62]</sup>。Jia 等<sup>[55]</sup>将降雨因素整合进 LSTM 网络(R-LSTM)。对于时间间隔为 10min 的预测,LSTM 和 R-LSTM 的预测准确率分别为 87.16% 和 88.31%,可以看出 R-LSTM 模型的准确率较高。在利用深度学习进行短时交通量预测时,应该考虑将更多的影响交通流的相关因素作为数据源,比如气象数据、道路基础设施灾病害信息等,这样能够提高模型的预测准确率和泛化性能。另外,该文献对比了 DBN 与 LSTM 在交通量预测中的表现,从实验结果可以看出 LSTM 的预测性能优于 DBN,这表明在交通量这样的时序数据预测中,LSTM 具有更好的预测性能。

门控循环单元神经网络(GRU)也是 RNN 的一种。与

LSTM 一样,GRU 网络也是为了解决长期依赖问题而提出的<sup>[63]</sup>;但与 LSTM 网络不同的是,GRU 网络将门控信号减少到了两个,这样可以在不影响模型训练效果的情况下提高训练效率,降低计算复杂度。

文献<sup>[56]</sup>对比了 LSTM 网络和 GRU 网络在交通量预测中的应用,二者的均方误差分别为 710.0502 和 668.9304,平均绝对百分比误差分别为 18.127758 和 17.2116。从两种误差评价指标来看,GRU 模型的预测性能优于 LSTM 模型,并且此种优势占据了整个时间序列的 84%。此外,GRU 模型的收敛速度快于 LSTM 模型,与理论分析相符。

RNN 通过门控机制可以学习到交通量数据的特征和长期依赖关系。现有研究在门控 RNN 能够对时间序列进行准确预测的基础上,尝试引入交通量的空间特征,以获取其时空

关联特性作为模型的输入,进而提高预测准确率。

### 3.2.2 卷积神经网络

卷积神经网络(CNN)<sup>[64]</sup>是一种前向传播算法。经典的卷积神经网络由卷积层、池化层和全连接层组成。卷积层通过稀疏连接、参数共享和等变表示3个重要的特点来提升算法能力,通过卷积运算可以增强原始参数特征并降低噪声数据带来的影响<sup>[65]</sup>。在池化层中,池化函数利用某位置相邻输出的总体统计特征来代替该位置的输出。由于池化引入了不变性,当输入经过少量的平移或旋转时,经过池化函数的大部分输出并不会发生改变,并且通过下采样可以减少计算量。全连接层将学习到的特征映射到样本标记空间。上述特征使得CNN网络能够有效应用于交通量预测中。

CNN是一种专门用来处理具有网络结构数据的神经网络。文献[66]将原始交通量数据处理成带有时空特征的交通量矩阵,这可以认为是在时间轴和空间轴上采样形成的二维网络数据。该文献采用了CNN和SVR的混合模型(CNN-SVR),其平均绝对误差相较于CNN模型减小了28%。CNN-SVR模型具有较高的预测准确率,并且受环境干扰较小,是

一种有效的短时交通量预测模型。

文献[67]提出的基于深度神经网络的交通量预测模型(DNN-BTF),由注意力模型、CNN、GRU及回归层组成。其中,注意力模型根据历史交通量、各观测点交通量与待预测观测点交通量的相关性,自动决定输入的权重,这样能够更为精准地利用交通量时空关联的特性,避免了粗略地利用先验知识来分配权重。考虑到交通量具有周期性的特点,分别为短时间内、前一日和前一周同时间点的交通量分配权重,然后将其作为输入,进而分别利用CNN和GRU来挖掘交通量的空间特征和时间特征,最后利用回归层对交通量进行预测。该模型在预测中充分利用了交通的时空关联特性,并且利用注意力模型计算权重,衡量了不同时间点、观测点的交通量输入对预测结果重要性的影响,为后续模型对时空特征的学习提供了准确的输入。

针对单变量交通数据时空特征的提取和多元交通数据相互依赖的特性,文献[68]提出了混合多模态深度学习框架(HMDLF),该模型由作为卷积模型的一维CNN和GRU模型及联合模型组成,其结构如图9所示。

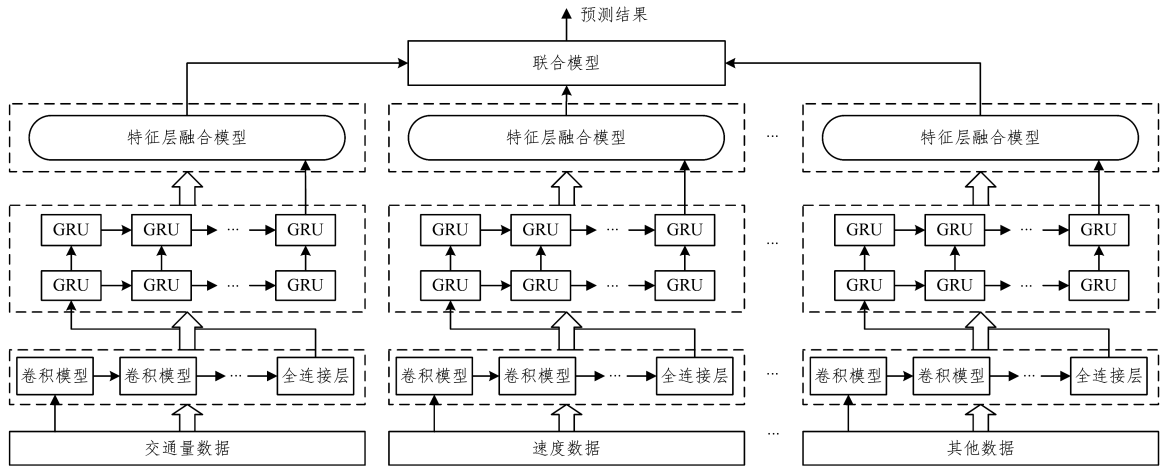


图9 混合多模态深度学习框架

Fig. 9 Hybrid multimodal deep learning framework

HMDLF基于多模态深度学习理论<sup>[69]</sup>,能够结合卷积模型和GRU模型的优势,并且能融合交通量、速度、出行时间等多源数据各自的优势。其中,卷积模型能够捕捉到局部交通数据变化趋势的空间特征,GRU模型能够捕捉到交通数据的短期时变和长期依赖的时间特征。继而其通过特征层融合得到各交通数据的时空特征,将这些时空特征联合作为回归层的输入,得到最终的交通量预测结果。该模型利用了多模型和多源数据的组合,能够降低噪声数据及随机因素对模型的影响,提升模型的预测准确率、泛化能力和鲁棒性。

由于CNN具有捕捉空间信息的特性,可利用它来提取交通量数据的空间特征。从现有研究成果可以看到,CNN结合门控RNN可以获取到交通量数据的时空关联特性,能够捕捉到短时交通量的时空特性及波动性的改变。

## 4 模型性能评价指标

性能评价指标是用来衡量模型好坏的重要标准,能够更好地分析模型的预测效果。在基于深度学习的短时交通量预

测研究中,常用的性能评价指标有平均绝对误差(MAE)、均方根误差(RMSE)、平均绝对百分误差(MAPE)和平均准确率(MA)。

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_i - y_i| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (2)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{x_i - y_i}{x_i} \right| \quad (3)$$

$$MA = 1 - MAPE \quad (4)$$

其中, $x_i$ 和 $y_i$ 分别代表在时间间隔 $i$ 的交通量实际观测值和预测值, $N$ 代表测试样本的数量。

MAE是预测值与真实值之间的平均绝对差异,能够有效地反映预测值误差的实际情况;RMSE用来反映预测值与真实值之间的偏差,能敏感地反映预测值中的极端误差,可以评价数据的变化程度,能够反映预测的精确度。MAPE能够衡量模型的预测准确率。表1按照时间顺序列出了基于深度学习的短时交通量预测中具有代表性的一些方法的性能及特点。

表1 基于深度学习的短时交通量预测代表性方法的性能对比

Table 1 Comparison of representative methods of short-term traffic flow forecasting based on deep learning

代表性方法	年份	MAE	RMSE	准确率/%	特点
Huang <sup>[26]</sup>	2014			91.74	多任务学习
Lv <sup>[37]</sup>	2015	34.1	50.0	93.25	堆栈自编码器,特征提取
Tan <sup>[32]</sup>	2016	22.26	30.82	90.79	算法优化
Koesdwiady <sup>[30]</sup>	2016	0.0405	0.0603		数据融合,多源数据,路网相关性分析
Luo <sup>[31]</sup>	2017			94.04	数据预处理,支持向量回归
Yang <sup>[39]</sup>	2017			90.37	数据预处理,模型训练,参数选择
Zhou <sup>[40]</sup>	2017		210.43	89.20	集成学习
Luo <sup>[66]</sup>	2017	8.2	10.0		数据预处理,模型训练,CNN
Jia <sup>[55]</sup>	2017	166.17	240.98	88.31	多源数据,对比 LSTM 和 DBN 预测性能
Zhao <sup>[54]</sup>	2017			93.78	路网相关性分析,特征提取
Fu <sup>[56]</sup>	2017	17.2116	25.8637		对比 LSTM 和 GRU 预测性能
Wu <sup>[67]</sup>	2018	22.1413	32.1639	92.13	注意力模型,时空关联分析,交通流周期性
Du <sup>[68]</sup>	2018		4.52		多源数据,多模态深度学习,特征提取

由于各文献的实验中采用了不同的数据采样时间间隔,因此 MAE 与 RMSE 两个评价指标具有不同的单位。对于上述两个评价指标,文献[32,37,56,66-67]使用的单位为 veh/5min,文献[68]使用的单位为 veh/15min,文献[40,55]使用的单位为 veh/h,文献[30]采用标准化后的数据进行性能评估。

## 5 结论和展望

利用深度学习进行短时交通量预测处于起步阶段,基于深度学习的短时交通量预测算法较传统方法有更高的预测准确率,鲁棒性更好,能捕捉到交通量数据的特征。可以预期,对于交通量预测问题,深度学习是一种具有前景的方法,值得推广和研究。以下总结了该研究在未来可能的发展趋势。

1)在如今有大量交通数据可供研究的背景下,预测模型的训练速度较慢,现有研究中无法充分利用这些数据来对模型进行训练。如何在保证训练精度的前提下,提升深度神经网络训练算法的效率,缩短训练时间,是一个有待解决的问题。另外,对于海量数据的处理,可以使用适用于深度学习的大规模数据处理并行框架,如 Spark<sup>[70-71]</sup>,或者使用可将矢量运算并行化执行的 GPU<sup>[72]</sup>及 FPGA 架构<sup>[73]</sup>。

2)深度模型超参数的学习。训练神经网络时,通常仅能对基础参数进行学习,而模型的超参数只能依靠人工根据经验设定,或者采取暴力枚举的尝试法来获取最优的超参数,这样会导致模型无法拟合现实问题的真实模型,易造成模型欠拟合或过拟合等问题。如何设计能够学习到模型超参数的算法,应成为深度学习的一个重要研究课题。

3)现有的预测模型大多仅考虑到交通量数据本身,忽略了其他交通特性对交通量预测的影响,较少的特征维度对模型的预测能力有较大的负面影响。因此,应将更多的交通特性融合进预测模型,考虑更多能够体现现实问题的数据特征,从而进行较为精准的建模。

4)现有的预测模型的种类较为单一,仍有很多适合于解决短时交通量预测的模型有待研究。例如,适合于处理序列数据的 RNN 具有多种变种,且每一种变种都能从不同角度来提升模型性能,如 Differential RNN 能够辨别出序列中的信息是否显著,并更好地捕捉到数据的时空模式<sup>[74]</sup>;Multidimensional LSTM 能够利用之前单元状态的联系来扩展 LSTM 的记忆<sup>[75]</sup>。这些模型都具有提高短时交通量预测性能的潜在能力。

5)由于道路交通环境具有复杂性及多样性的特点,因此如何提升深度模型的推广能力和泛化性能,是利用深度学习进行短时交通量预测需要着重考虑的一个问题。

**结束语** 本文对目前基于深度学习的短时交通量预测方法进行了概述与总结,并指出了现有研究中尚存的问题和不足,总结了该研究在未来可能的发展趋势。虽然,深度学习在短时交通量预测的应用中还不成熟,但就已有研究所取得的成果来看,深度学习预测方法已较传统预测方法有了较大的提升,但仍然具有很大的改进提升空间。可以预期,在传统短时交通量预测方法遇到瓶颈的情况下,利用深度学习方法对短时交通量进行预测将会成为该领域未来的研究热点。

## 参考文献

- [1] HINTON G E, OSINDERO S, TEH Y W. A Fast Learning Algorithm for Deep Belief Nets[J]. *Neural Computation*, 2006, 18(7):1527-1554.
- [2] AHMED M S, COOK A R. Analysis of Freeway Traffic Time-Series Data by Using Box-Jenkins Techniques[J]. *Transportation Research Record*, 1979, 722:1-9.
- [3] LEVIN M, TSAO Y D. On Forecasting Freeway Occupancies and Volumes[J]. *Transportation Research Record*, 1980(722):47-49.
- [4] WILLIAMS B M, HOEL L A. Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results[J]. *Journal of Transportation Engineering*, 2003, 129(6):664-672.
- [5] LEE S, FAMBRO D. Application of Subset Autoregressive Integrated Moving Average Model for Short-Term Freeway Traffic Volume Forecasting[J]. *Transportation Research Record Journal of the Transportation Research Board*, 1999, 1678(1):179-188.
- [6] OKUTANI I, STEPANEDES Y J. Dynamic Prediction of Traffic Volume through Kalman Filtering Theory[J]. *Transportation Research Part B*, 1984, 18(1):1-11.
- [7] GUO J, HUANG W, WILLIAMS B M. Adaptive Kalman Filter Approach for Stochastic Short-Term Traffic Flow Rate Prediction and Uncertainty Quantification[J]. *Transportation Research Part C*, 2014, 43:50-64.
- [8] ZHANG Y, XIE Y. Forecasting of Short-Term Freeway Volume with v Support Vector Machines[J]. *Transportation Research Record Journal of the Transportation Research Board*, 2007, 2024(1):92-99.
- [9] DAVIS G A, NIHAN N L. Nonparametric Regression and Short-Term Freeway Traffic Forecasting[J]. *Journal of Transportation Engineering*, 1991, 117(2):178-188.
- [10] SMITH B L, DEMETSKY M J. Traffic Flow Forecasting; Com-

- parison of Modeling Approaches[J]. *Journal of Transportation Engineering*, 1997, 123(4): 261-266.
- [11] WU C H, HO J M, LEE D T. Travel-Time Prediction with Support Vector Regression [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2004, 5(4): 276-281.
- [12] XIE Y, ZHAO K, SUN Y, et al. Gaussian Processes for Short-Term Traffic Volume Forecasting [J]. *Transportation Research Record Journal of the Transportation Research Board*, 2010, 2165: 69-78.
- [13] ZHANG L, LIU Q, YANG W, et al. An Improved K-Nearest Neighbor Model for Short-Term Traffic Flow Prediction [J]. *Procedia-Social and Behavioral Sciences*, 2013, 96: 653-662.
- [14] ZHENG Z, SU D. Short-Term Traffic Volume Forecasting: A K-Nearest Neighbor Approach Enhanced by Constrained Linearly Sewing Principle Component Algorithm [J]. *Transportation Research Part C Emerging Technologies*, 2014, 43: 143-157.
- [15] WANG J, DENG W, ZHAO J B. Short-Term Freeway Traffic Flow Prediction Based on Improved Bayesian Combined Model [J]. *Journal of Southeast University (Natural Science Edition)*, 2012, 42(1): 162-167. (in Chinese)  
王建, 邓卫, 赵金宝. 基于改进型贝叶斯组合模型的短时交通流量预测[J]. *东南大学学报(自然科学版)*, 2012, 42(1): 162-167.
- [16] SMITH B L, DEMETSKY M J. Short-Term Traffic Flow Prediction: Neural Network Approach [J]. *Transportation Research Record*, 1994(1453): 98-104.
- [17] MESSER C, THOMAS URBANIK I I. Short-Term Freeway Traffic Volume Forecasting Using Radial Basis Function Neural Network [J]. *Transportation Research Record Journal of the Transportation Research Board*, 1998(1651): 39-47.
- [18] BENGIO Y, LECUN Y. Scaling Learning Algorithms Towards AI [J]. *Large-Scale Kernel Machines*, 2007, 34(5): 1-41.
- [19] ERFANI S M, RAJASEGARAR S, KARUNASEKERA S, et al. High-Dimensional and Large-Scale Anomaly Detection Using a Linear One-Class SVM with Deep Learning [J]. *Pattern Recognition*, 2016, 58(C): 121-134.
- [20] LECUN Y, BENGIO Y, HINTON G. Deep Learning [J]. *Nature*, 2015, 521(7553): 436-444.
- [21] BENGIO Y. Learning Deep Architectures for AI [J]. *Foundations & Trends in Machine Learning*, 2009, 2(1): 1-127.
- [22] LI D, DONG Y. Deep Learning: Methods and Applications [J]. *Foundations & Trends® in Signal Processing*, 2014, 7(3): 197-387.
- [23] SUN Z Y, LU C X, SHI Z Z, et al. Research and Advances on Deep Learning [J]. *Computer Science*, 2016, 43(2): 1-8. (in Chinese)  
孙志远, 鲁成祥, 史忠植, 等. 深度学习研究与进展 [J]. *计算机科学*, 2016, 43(2): 1-8.
- [24] JIA J P, QIN Y H. Survey on Visual Tracking Algorithms Based on Deep Learning Tehnologies [J]. *Computer Science*, 2017, 44(s1): 19-23. (in Chinese)  
贾静平, 覃亦华. 基于深度学习的视觉跟踪算法研究综述 [J]. *计算机科学*, 2017, 44(s1): 19-23.
- [25] ROUX N L, BENGIO Y. Representational Power of Restricted Boltzmann Machines and Deep Belief Networks [J]. *Neural Computation*, 2008, 20(6): 1631-1649.
- [26] HUANG W, SONG G, HONG H, et al. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks with Multitask Learning [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(5): 2191-2201.
- [27] CARUANA R. Multitask Learning [J]. *Machine Learning*, 1997, 28(1): 41-75.
- [28] SOUA R, KOESDWIADY A, KARRAY F. Big-Data-Generated Traffic Flow Prediction Using Deep Learning and Dempster-Shafer Theory [C] // *International Joint Conference on Neural Networks*. 2016: 3195-3202.
- [29] DEMPSTER A P. The Dempster-Shafer Calculus for Statisticians [J]. *International Journal of Approximate Reasoning*, 2007, 48(2): 365-377.
- [30] KOESDWIADY A, SOUA R, KARRAY F. Improving Traffic Flow Prediction with Weather Information in Connected Cars: A Deep Learning Approach [J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(12): 9508-9517.
- [31] LUO X L, JIAO Q Q, NIU L Y, et al. Short-Term Traffic Flow Prediction Based on Deep Learning [J]. *Application Research of Computers*, 2017, 34(1): 91-93. (in Chinese)  
罗向龙, 焦琴琴, 牛力瑶, 等. 基于深度学习的短时交通流预测 [J]. *计算机应用研究*, 2017, 34(1): 91-93.
- [32] TAN H, XUAN X, WU Y, et al. A Comparison of Traffic Flow Prediction Methods Based on DBN [C] // *Cota International Conference of Transportation Professionals*. Shanghai, China, 2016: 273-283.
- [33] HINTON G E. A Practical Guide to Training Restricted Boltzmann Machines [J]. *Momentum*, 2012, 9(1): 599-619.
- [34] HINTON G E, SALAKHUTDINOV R. Using Deep Belief Nets to Learn Covariance Kernels for Gaussian Processes [C] // *International Conference on Neural Information Processing Systems*. 2008: 1249-1256.
- [35] HINTON G E, SALAKHUTDINOV R. Reducing the Dimensionality of Data with Neural Networks [J]. *Science*, 2006, 313(5786): 504-507.
- [36] PALM R B. Prediction as a Candidate for Learning Deep Hierarchical Models of Data [D]. *Kongens Lyngby: Technical University of Denmark*, 2012.
- [37] 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.
- [38] DUAN Y, LV Y, WANG F. Performance Evaluation of the Deep Learning Approach for Traffic Flow Prediction at Different Times [C] // *International conference on service operations and logistics, and informatics*. IEEE, 2016: 223-227.
- [39] YANG H F, DILLON T S, CHEN Y P. Optimized Structure of the Traffic Flow Forecasting Model with a Deep Learning Approach [J]. *IEEE transactions on neural networks and learning systems*, 2017, 28(10): 2371-2381.
- [40] ZHOU T, HAN G, XU X, et al.  $\delta$ -Agree Adaboost Stacked Autoencoder for Short-Term Traffic Flow Forecasting [J]. *Neurocomputing*, 2017, 247: 31-38.
- [41] TANG J, XU G, WANG Y, et al. Traffic Flow Prediction Based on Hybrid Model Using Double Exponential Smoothing and Support Vector Machine [C] // *International IEEE Conference on Intelligent Transportation Systems*. 2014: 130-135.
- [42] KANZOW C, YAMASHITA N, FUKUSHIMA M. Levenberg-Marquardt Methods with Strong Local Convergence Properties for Solving Nonlinear Equations with Convex Constraints [M].

- Elsevier Science Publishers B V, 2004; 321-343.
- [43] TSAI J T, CHANG C C, CHEN W P, et al. Optimal Parameter Design for Ic Wire Bonding Process by Using Fuzzy Logic and Taguchi Method[J]. *IEEE Access*, 2017, 4: 3034-3045.
- [44] KAWAGUCHI K. Deep Learning without Poor Local Minima [C]// the 30th Conference on Neural Information Processing Systems, Barcelona, Spain, 2016; 586-594.
- [45] SCHAPIRE R E, FREUND Y. Boosting: Foundations and Algorithms[J]. *Kybernetes*, 2013, 42(1): 164-166.
- [46] PINEDA F J. Generalization of Back-Propagation to Recurrent Neural Networks[J]. *Physical Review Letters*, 1987, 59(19): 2229-2232.
- [47] SUTSKEVER I, MARTENS J, HINTON G E. Generating Text with Recurrent Neural Networks[C]// International Conference on Machine Learning, Bellevue, Washington, USA, 2011; 1017-1024.
- [48] ELMAN J L. Finding Structure in Time[J]. *Cognitive Science*, 1990, 14(2): 179-211.
- [49] VAN LINT J, HOOGENDOORN S, VAN ZUYLEN H. Freeway Travel Time Prediction with State-Space Neural Networks: Modeling State-Space Dynamics with Recurrent Neural Networks[J]. *Transportation Research Record: Journal of the Transportation Research Board*, 2002(1811): 30-39.
- [50] BENGIO Y, SIMARD P, FRASCONI P. Learning Long-Term Dependencies with Gradient Descent Is Difficult [J]. *IEEE Transactions on Neural Networks*, 1994, 5(2): 157-166.
- [51] PASCANU R, MIKOLOV T, BENGIO Y. On the Difficulty of Training Recurrent Neural Networks[C]// International Conference on Machine Learning, Atlanta, GA, USA, 2013; 1310-1318.
- [52] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [53] LE Q V, JAITLY N, HINTON G E. A Simple Way to Initialize Recurrent Networks of Rectified Linear Units[J]. *arXiv*: 1504.00941, 2015.
- [54] ZHAO Z, CHEN W, WU X, et al. LSTM Network: A Deep Learning Approach for Short-Term Traffic Forecast[J]. *IET Intelligent Transport Systems*, 2017, 11(2): 68-75.
- [55] JIA Y, WU J, XU M. Traffic Flow Prediction with Rainfall Impact Using a Deep Learning Method[J]. *Journal of Advanced Transportation*, 2017, 2017(722): 1-10.
- [56] FU R, ZHANG Z, LI L. Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction[C]// Youth Academic Annual Conference of Chinese Association of Automation (YAC). Wuhan, China, 2017; 324-328.
- [57] TIAN Y, PAN L. Predicting Short-Term Traffic Flow by Long Short-Term Memory Recurrent Neural Network[C]// IEEE International Conference on Smart City. Chengdu, China, 2015; 153-158.
- [58] CHEN Y Y, LV Y S, LI Z J, et al. Long Short-Term Memory Model for Traffic Congestion Prediction with Online Open Data [C]// IEEE International Conference on Intelligent Transportation Systems. Rio de Janeiro, Brazil, 2016; 132-137.
- [59] SHAO H, SOONG B H. Traffic Flow Prediction with Long Short-Term Memory Networks (LSTMs)[C]// Region 10 Conference (TENCON). IEEE, 2017; 2986-2989.
- [60] XUE W X, XU L H. Short-Term Traffic Flow Prediction Based on Deep Learning[J]. *Journal of Transportation Systems Engineering and Information Technology*, 2018, 18(1): 81-88. (in Chinese)
- 王祥雪, 许论辉. 基于深度学习的短时交通流预测研究[J]. *交通运输系统工程与信息*, 2018, 18(1): 81-88.
- [61] CHUNG E. Does Weather Affect Highway Capacity [C]// 5th International Symposium on Highway Capacity and Quality of Service. Yakoma, Japan, 2006.
- [62] MAZE T H, AGARWAI M, BURCHETT G. Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow[J]. *Transportation Research Record Journal of the Transportation Research Board*, 2006, 1948(1): 170-176.
- [63] TANG Y, HUANG Y, WU Z, et al. Question Detection from Acoustic Features Using Recurrent Neural Network with Gated Recurrent Unit[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2016; 6125-6129.
- [64] RAWAT W, WANG Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review[J]. *Neural Computation*, 2017, 29(9): 2352-2449.
- [65] GOODFELLOW I, BENGIO Y, COURVILLE A. Deep Learning [M]. Cambridge, Massachusetts, USA: MIT Press, 2016.
- [66] LUO W H, DONG B T, WANG Z S. Short-Term Traffic Flow Prediction Based on CNN-SVR Hybrid Deep Learning Model [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2017, 17(5): 68-74. (in Chinese)
- 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.
- [67] WU Y K, TAN H C, QIN L Q, et al. A Hybrid Deep Learning Based Traffic Flow Prediction Method and Its Understanding [J]. *Transportation Research Part C: Emerging Technologies*, 2018, 90: 166-180.
- [68] DU S, LI T, GONG X, et al. A Hybrid Method for Traffic Flow Forecasting Using Multimodal Deep Learning[J]. *arXiv*: 1803.02099, 2018.
- [69] NGIAM J, KHOSLA A, KIM M, et al. Multimodal Deep Learning[C]// International Conference on Machine Learning. Bellevue, Washington, USA, 2011; 689-696.
- [70] SHANAHAN J, LIANG D. Large Scale Distributed Data Science from Scratch Using Apache Spark 2.0[C]// International Conference on World Wide Web Companion. New York: ACM, 2017; 955-957.
- [71] ALSHEIKH M A, NIYATO D, LIN S, et al. Mobile Big Data Analytics Using Deep Learning and Apache Spark [J]. *IEEE Network*, 2016, 30(3): 22-29.
- [72] COELHO I M, COELHO V N, LUZ E J D S, et al. A GPU Deep Learning Metaheuristic Based Model for Time Series Forecasting[J]. *Applied Energy*, 2017, 201: 412-418.
- [73] WANG C, GONG L, YU Q, et al. Dlau: A Scalable Deep Learning Accelerator Unit on FPGA [J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2017, 36(3): 513-517.
- [74] VEERIAH V, ZHUANG N, QI G J. Differential Recurrent Neural Networks for Action Recognition [C]// IEEE International Conference on Computer Vision. IEEE, 2015; 4041-4049.
- [75] GRAVES A, FERNÁNDEZ S, SCHMIDHUBER J. Multi-Dimensional Recurrent Neural Networks[C]// International Conference on Artificial Neural Networks. Berlin: Springer, 2007; 549-558.