



北京大学

# 本科毕业论文

题目： 基于城市路网和神经网络的  
实时交通速度预测

姓 名： 步凡

学 号： 1300010677

院 系： 数学科学学院

专 业： 数据科学与大数据技术

导师姓名： 李铁军 教授

二〇一七年六月

## 版权声明

任何收存和保管本论文各种版本的单位和个人，未经本论文作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则，引起有碍作者著作权之问题，将可能承担法律责任。

## 摘要

基于实时数据的交通路况预测技术是近年来的热门研究课题之一。准确高效的路段交通速度预测，因其应用广泛、影响积极，故而受到研究界的格外重视，而能够处理大规模、复杂城市路网结构下的速度预测问题的模型，则亟待发展。

本文从人工神经网络的最新研究进展得到启发，将图上的卷积神经网络与卷积长短期记忆网络相结合，建立图上的卷积长短期记忆网络 GConvLSTM，能够在对交通速度的时间序列属性进行编码的同时，将道路网络的空间结构属性也融合到模型之中，从而对指定区域内全部道路的交通速度进行批量、高效、准确的预测。

在真实数据集上的实验表明，GConvLSTM 模型不仅在预测准确率上优于基准模型，在速度变化趋势的预测上也较为敏锐，并且预测结果具有较小的噪声水平。

**关键词：**交通速度预测，卷积神经网络，长短期记忆网络

# Traffic Speed Nowcasting Based on Urban Road Network and Artificial Neural Network

Fan Bu (Data Science and Big Data Technology)

Directed by Prof. Tiejun Li

## ABSTRACT

Traffic forecasting technology based on actual-time data is one of the most researched topics of recent years. Accurate traffic speed prediction is deemed particularly important by the academia due to the large variety of its potential applications and positive influence. However, the models capable of tackling traffic speed forecasting at a large scale and on complex urban road networks are still under-developed.

This paper gets inspiration from the latest progress in the field of artificial neural networks, integrates CNN on graphs and Convolutional LSTM, and develops the novel model Graph Convolutional LSTM, which is able to encode both the temporal and spatial factors in traffic speed information on transportation networks and thus produce traffic speed forecasting on all the roads in a given district collectively, efficiently, and accurately.

Experiments on a real dataset demonstrate that the model not only outperforms baselines in accuracy, but also detects the change of speed tendency more sharply and provides prediction results with a lower noise level.

**KEY WORDS:** Traffic Speed Forecasting, CNN, LSTM

## 目录

第一部分	引言 .....	1
第二部分	论文主体 .....	2
1	问题描述 .....	2
2	模型构建 .....	2
2.1	路网结构的表示 .....	2
2.2	基于 GRAPH CNN 对路网结构图上的信号编码 .....	3
2.3	基于 CONVLSTM 进行历史数据编码和数据预测 .....	3
2.4	综合 GRAPH CNN 和 CONVLSTM 的 GCONVLSTM 模型 .....	4
3	数据实验 .....	5
3.1	数据描述与数据预处理 .....	5
3.2	超参数设置与实验细节 .....	6
3.3	实验结果 .....	7
4	结论与展望 .....	10
参考文献	.....	12
致谢	.....	14
北京大学学位论文原创性声明和使用授权说明	.....	15

## 第一部分 引言

实时交通路况预测是近年来愈发成为热门的研究课题之一。随着交通路况数据采集技术的成熟与丰富以及大规模数据存储与处理技术的高速发展,如今,城市道路上的实时交通数据的获取和存储已不再是问题,研究的重点已经转向这类数据的处理、分析与利用,而其中的一个主要方向即是基于实时交通数据的路况预测。

对路段交通流量或速度的高效、准确的预测有着多方面的意义:出行的市民可以据此选择交通工具、调整出行时间,导航或打车软件的开发者可以据此优化其路线规划或估计到达时间的算法,而城市交通管理部门可以据此发布拥堵预警、调节路口信号灯和潮汐车道以保障城市交通顺畅运行。

另一方面,自上世纪末以来,人工神经网络逐渐成为应用科学界的重要研究工具,近十年来,其发展可谓如火如荼,在图像和视频分析、自然语言处理等领域,各种结构的人工网络已经成为占据统治地位的标准模型。不难想象,在交通路况预测甚至更为广泛的交通问题研究中,人工神经网络也将拥有较为显著的一席之地。

在本文的涉及的研究领域,前人主要进行了以下两方面的研究工作:一,在交通预测领域,使用包括神经网络在内的各种模型,对一条或多条道路上的交通速度或流量进行中长时预测;二,在人工神经网络领域,使用循环神经网络进行视频下一帧或动态图像的预测,以及建立对图等非对称结构编码的卷积神经网络。

在交通预测领域,较为有效的传统预测模型主要包括时间序列模型和贝叶斯网络等传统机器学习模型,在精细的模型设置下,传统模型也可以较好地融合时间和空间两方面信息,做出较为准确的速度(或流量)预测。文献[1]构建了基于混合高斯模型(GMM)的贝叶斯网络(Bayesian Network),使用贝叶斯网络结构中的某些边来表示临近路段对目标路段的交通路况的影响,并使用在北京市东三环附近的一个有10个路口的小型区域上采集的流量数据进行实验;文献[2]则是众多时间序列预测模型中成果较新、表现较好的一篇,与前人通常使用的AR(自回归)模型、ARIMA(差分自回归滑动平均)模型或季节ARIMA不同的是,该文献建立了“多元空间-时间自回归”(MSTAR)模型,在向量自回归滑动平均模型的基础上加入路段间的空间连接关系,并使用在16个路段组成的有一个交汇点的样本路网上采集的速度数据进行实验。近几年来,神经网络,特别是深度网络,因其捕捉非线性特征和处理大规模数据的能力,已经越来越被广泛地应用于交通速度预测。文献[3]使用基于自动编码器(Auto Encoder)的深度网络进行交通流量的预测,并在芝加哥I-55高速公路上采集的数据上进行实验;文献[4]建立了堆栈式自动编码器(Stacked Auto Encoder)的深度网络,使用在加州全境高速公路上采集的交通流量数据进行实验;文献[5]使用长短期记忆网络(LSTM Network)进行交通速度预测,并在北京市二环路主路德胜门到安定门路段上采集的速度数据进行实验;文献[6]使用深度信念网络(DBN)进行交通速度预测,并在北京市德胜门外大街主路德胜门到马甸桥路段上采集的数据上进行实验。

综合前人在交通预测领域进行的尝试,不难发现,尽管模型的复杂度和预测准确性在不断增加,但这些模型普遍侧重于交通预测的时间属性,而轻视甚至忽略了空间属性:文献[1]和文献[2]虽然有意识地将道路路网结构融入模型中,但模型的基础构建直接受到路网结构的影响,并且模型所能应用的路网结构简单、规模较小;文献[3]、[5]、[6]的模型只能应用在仅具前后继关系的单一道路或者无空间结构关系的单一路段上,而文献[4]尽管可以处理一个较大地区的路网,却回避了处理路网结构的问题,只将全部道路的数据直接输入网络,并且也只能处理高速公路,并非全路网。

在人工神经网络领域,视频预测是较为热门的研究课题,所使用的模型也比较多样。文献[7]

对传统的长短期记忆网络改进为卷积长短期记忆网络，使网络更好地捕捉每一帧图像的局部信息，从而得到更精确的预测，在雷达图像降水预测中得到了很好的结果；文献 [8] 建立的深度网络通过在每一层学习上一层的预测误差来增进网络预测下一帧图像的精确度；文献 [9] 构建了多卷积层与长短期记忆层的深度网络，在预测视频下一帧的同时能够更好地把握视频中主要物体的动态，在计算机视觉领域有很好的应用前景。在这些工作中，视频的每一帧均为图像，即排列在结构高度对阵的长方形格点上的像素点数值，而当每一帧信号的内部结构为更加复杂或不对称的图结构时，每一帧信号就无法用传统卷积神经网络（CNN）的方式编码。近年来，一些学者致力于研究图上的卷积运算，根据文献 [10]，在图上的卷积神经网络（Graph CNN）大致有两种实现方式：第一，直观地定义图中每个顶点的邻域，在其邻域顶点中进行卷积运算，如文献 [11]；第二，基于图的谱分析，将卷积运算化为卷积核信号与图的拉普拉斯矩阵间的矩阵运算，如文献 [12]。图上的卷积神经网络的构建使图结构的信号可以编码作为传统深度网络的输入信号，并且保留卷积神经网络的局部化、权重共享、参数较少等优良性质。

本文工作的目标即是将人工神经网络领域的最新进展与交通预测领域结合起来，使用图上的卷积神经网络编码路网结构信息，再使用能够进行视频预测的网络编码该信息的时间序列，构建出同时处理城市路网的空间拓扑结构和交通速度的时间序列属性的预测模型，使模型具有较高预测准确性的同时，能够处理较复杂、较大规模的路网，甚至能够拓展到整个城市的全路网上。

在接下来的论文主体部分中，主要包括以下几个章节：第 1 小节对问题进行更为详细的阐述并引入一些重要的基础概念，第 2 小节介绍模型的各部分结构和模型整体算法，第 3 小节给出真实数据集上的实验结果并作出分析，第 4 小节进行总结并对未来的工作做出展望。

## 第二部分 论文主体

### 1 问题描述

本文解决的问题是，对已知路网结构的若干给定路段  $1, 2, \dots, R$  和这些路段上的历史速度数据  $\{\mathbf{x}_{t-L+1}, \mathbf{x}_{t-L+2}, \dots, \mathbf{x}_{t-1}, \mathbf{x}_t\}$  ( $\mathbf{x}_t = [x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(R)}]^T$ ，其中  $x_t^{(r)}$  为路段  $r$  在时间  $t$  的速度)，给出时间  $T$  后这些路段上的速度  $\mathbf{x}_{t+T}$  的预测  $\hat{\mathbf{x}}_{t+T}$ 。

在本文的情景中，使用的速度皆为一定时间间隔内（通常是 5 分钟内）的路段平均速度， $L$  通常取为 12（即使用长度为 1 小时的历史）， $T$  通常取为 6（即预测 30 分钟后的路段速度）。非常自然地，我们可以假设，在远长于预测所用的历史和预测时长的时间内，目标路段的内部路网结构不变。

评价预测误差的指标选为 MAPE (Mean Absolute Percentage Error)，其定义为

$$MAPE_t = \frac{1}{R} \sum_{r=1}^R \frac{|\hat{x}_t^{(r)} - x_t^{(r)}|}{x_t^{(r)}}$$

### 2 模型构建

#### 2.1 路网结构的表示

本文采用图 (Graph) 表示实际路网结构。与直观上的以道路交叉口或连接处为顶点、路段为边的地图式路网不同的是，这里使用以路段为顶点、路段交叉口或连接处为边的有向图。

以一个包含路段  $\{1, 2, \dots, R\}$  的目标区域为例，它的路网结构图  $G = \{V, E\}$  是这样的： $V = \{1, 2, \dots, R\}$ ，有向边  $(r_1, r_2) \in E$  当且仅当机动车可以从路段  $r_1$  直接行驶到路段  $r_2$ ， $G$  的邻接矩阵  $W = (w_{r_1 r_2})$  满足

$$w_{r_1 r_2} = \begin{cases} 1 & (r_1, r_2) \in E \\ 0 & \text{否则} \end{cases}$$

于是，在时间  $t$  的路段速度  $\mathbf{x}_t$  即可视为图  $G$  顶点集  $V$  上的一组信号。

## 2.2 基于 Graph CNN 对路网结构图上的信号编码

本文采用文献 [12] 提出的局部化谱滤波的方法定义图  $G = \{V, E, W\}$  上的卷积运算。简单来说，对图顶点上的一组信号  $x$  做卷积运算，可以转化为对其拉普拉斯矩阵  $L^1$  做运算，再作用在信号  $x$  上。设卷积核为函数  $g_\theta$ ，则经过  $g_\theta$  滤波后信号  $x$  变为

$$y = g_\theta(L)x.$$

若使该运算具有  $K$ -局部性（即只有与顶点  $i$  距离不超过  $K$  的顶点才能被包含在以顶点  $i$  为中心的卷积核中），则可设函数  $g_\theta$  为如下形式的多项式函数

$$g_\theta(L) = \sum_{k=0}^{K-1} \theta_k L^k.$$

$\theta \in \mathbb{R}^K$  为该多项式的参数向量。可以证明， $d_G(i, j) > K \Rightarrow (L^K)_{i,j} = 0$ ，其中  $d_G$  为图  $G$  上两顶点间最短距离。

在网络训练中，对  $F_{in}$  组输入信号经过卷积运算得到  $F_{out}$  组输出特征的计算式如下

$$y_j = \sum_{i=1}^{F_{in}} g_{\theta_{i,j}}(L)x_i,$$

其中  $i = 1, 2, \dots, F_{in}$ ,  $j = 1, 2, \dots, F_{out}$ ,  $x_i \in \mathbb{R}^{\#(V)}$ ,  $\theta_{i,j} \in \mathbb{R}^K$  为待训练的参数或已给定的卷积权重参数。

## 2.3 基于 ConvLSTM 进行历史数据编码和速度预测

本文使用文献 [7] 提出的名为 ConvLSTM 的卷积长短期记忆网络模型处理道路交通历史数据的时间序列信息并进行速度预测。文献 [7] 提出的模型对经典的全连接长短期记忆网络 (FC-LSTM) 进行改造，将其中的全连接部分改为卷积，这样可以更好地抓取信号的局部信息，并大大减少参数的数量。

具体来说，传统的 FC-LSTM 在时间步  $t$  对每一个网络单元各状态值的更新计算公式为

<sup>1</sup>这里  $L = W - D$ ,  $D$  为满足  $D_{ii} = \sum_j W_{ij}$  的对角矩阵。



$$\begin{aligned}
 i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \\
 f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \\
 c_t &= f_t \circ c_{t-1} + i_t \circ \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\
 o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \circ c_t + b_o) \\
 h_t &= o_t \circ \tanh(c_t)
 \end{aligned}$$

其中  $i_t, f_t, c_t, o_t, h_t$  分别为输入门限、遗忘门限、单元状态、输出门限和单元最终状态,  $x_t$  和  $c_t, h_t$  均为一维向量,  $\sigma()$  为 Sigmoid 函数, ‘ $\circ$ ’ 为 Hadamard 乘积。模型中待训练的参数为  $W$  和  $b$ 。文献 [7] 提出的 ConvLSTM 将上述运算公式中的普通乘积改为卷积运算, 即

$$\begin{aligned}
 i_t &= \sigma(W_{xi} * \mathcal{X}_t + W_{hi} * \mathcal{H}_{t-1} + W_{ci} \circ \mathcal{C}_{t-1} + b_i) \\
 f_t &= \sigma(W_{xf} * \mathcal{X}_t + W_{hf} * \mathcal{H}_{t-1} + W_{cf} \circ \mathcal{C}_{t-1} + b_f) \\
 \mathcal{C}_t &= f_t \circ \mathcal{C}_{t-1} + i_t \circ \tanh(W_{xc} * \mathcal{X}_t + W_{hc} * \mathcal{H}_{t-1} + b_c) \\
 o_t &= \sigma(W_{xo} * \mathcal{X}_t + W_{ho} * \mathcal{H}_{t-1} + W_{co} \circ \mathcal{C}_t + b_o) \\
 \mathcal{H}_t &= o_t \circ \tanh(\mathcal{C}_t)
 \end{aligned}$$

这里  $\mathcal{X}_t, \mathcal{C}_t, \mathcal{H}_t$  都为三维张量 (可以有某些维度的长度为 1), ‘ $*$ ’ 为卷积运算。需要说明的是, 这里  $\mathcal{X}_t$  是将每一帧的信号进行粗粒度化得到的三维张量, 比如对于一张  $64 \times 64$  像素的图像, 将每 16 个像素点“立起来”, 则成为一个  $16 \times 16 \times 16$  的张量信号。ConvLSTM 的单元, 和 FC-LSTM 一样, 可以堆叠起来组成多层网络用于编码历史序列。

在预测时, 文献 [7] 使用多层 ConvLSTM 单元构成的预测网络, 这些单元的各初始状态直接拷贝编码网络的最终单元状态, 在预测网络的最后一层, 将各层状态连接起来, 再通过一个卷积核大小为  $1 \times 1$  的卷积层得到预测结果。文献 [7] 使用图 2.1 解释其编码-预测网络的流程<sup>2</sup>: 编码网络将整个输入序列压缩为隐状态的张量, 而预测网络则将该张量展开以形成最终预测。

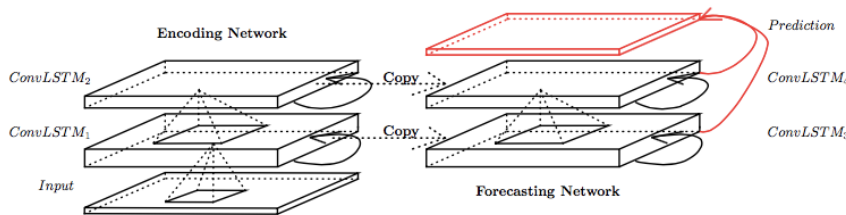


图 2.1: ConvLSTM 的编码-预测网络结构

## 2.4 综合 Graph CNN 和 ConvLSTM 的 GConvLSTM 模型

综合以上两部分, 本文提出可用于图上信号预测的 GConvLSTM 模型: 其编码和预测主体结构采用 ConvLSTM 的编码-预测网络结构, 模型训练过程中各单元状态更新采用 ConvLSTM 的计算公

<sup>2</sup>该图片直接截取自文献 [7]

式，但与输入信号相关的所有卷积运算都使用 4.2 介绍的方法（记为 `graph_conv` 运算），并且在预测网络最后一层，采用局部化系数  $K = 1$  的 `graph_conv` 运算的卷积层做出最终预测。

尽管 ConvLSTM 处理的输入序列每一帧为二维图像，本文情境下的路段交通速度序列同样可以作为输入信号：对于一个  $R \times 1$  大小的道路速度信号，如果无须进行粗粒度化，则也可将其当作一个  $R \times 1 \times 1$  的三维张量信号输入模型；如果有粗粒度化的需要，则对于  $R$  的某个正因子  $p$ ，该信号也可以化为  $R/p \times 1 \times p$  的三维张量，这样在  $R$  较大时，可以减少需要训练的模型参数数量。

在下面的算法 2.1 中，列出了 GConvLSTM 模型对历史数据进行编码和模型训练与测试的算法整体结构。

---

**Algorithm 2.1** GConvLSTM 模型的算法主体

---

- 1: 给定区域道路的路网结构图  $G$ ，局部化系数  $K$ ，训练样本  $X_{train}$ ，测试样本  $X_{test}$ ，网络各层通道数  $f_{out-list}$ ，网络输入层通道数  $f_{in}$ ，迭代次数  $M$ ，学习策略  $ls$ ，训练批量  $bs$
  - 2: 计算拉普拉斯矩阵  $L$ ，将训练样本  $X_{train}$  和测试样本  $X_{test}$  转化为三维张量的有序集  $\mathcal{X}_{train}$  和  $\mathcal{X}_{test}$
  - 3: 随机初始化网络参数； $m = 0$
  - 4: 若  $m < M$ ，从  $\mathcal{X}_{train}$  中随机选取大小为  $bs$  的样本序列集  $\mathcal{X}_{batch}$ ， $t = 1$ ；否则迭代结束，输出并保存结果
  - 5: 若  $t \leq 12$ ，使用  $\mathcal{X}_{batch}$  的时间步  $t$  的信号依次计算  $f_{in}$  个输入特征和以  $f_{out-list}$  长度为层数、各分量为个数的输出特征， $t \leftarrow t + 1$ ；若  $12 < t \leq 18$ ，使用时间步  $t - 1$  计算得到的各输出特征值连接起来，使用 `graph_conv` 卷积计算得到预测值，并与样本比较，计算 L2 损失和  $t = 18$  时的  $MAPE$ ， $t \leftarrow t + 1$ ；否则转步 6
  - 6: 使用学习策略  $ls$  优化时间步  $12 < t \leq 18$  的 L2 损失之和；若  $10|m$ ，对  $\mathcal{X}_{test}$  进行步 5 中的操作，并保存输出训练集和测试集上的 L2 损失和  $t = 18$  时的  $MAPE$ ； $m \leftarrow m + 1$ ，转步 4
- 

## 3 数据实验

### 3.1 数据描述与数据预处理

本文使用北京大学大数据研究院“交通大数据”课程中提供的交通速度数据集进行数据实验。该数据采集自北京市东四环机场高速入口至十八里店桥南北双方向主辅路共 12 个路段，采集时间为 2016 年 4 月 14 日至 2016 年 4 月 27 日（共 14 天），速度数据为每天从早 0 点到晚 24 点每 5 分钟间隔各路段上的机动车平均行驶速度（单位：km/h），即每天每路段有 288 条记录。

对这 12 个路段进行编号（主路路段以 1 开头，辅路路段以 2 开头），路段编号与路段描述的对对应关系见表 3.1。由于数据本身并不包含这些路段间的空间连接关系，故采用手工地图搜索的方式<sup>3</sup>获得这 12 个路段间的空间连接关系，从而获得其路网结构图  $G$ ，其大致结构如图 3.1 所示。

---

<sup>3</sup>这里使用了谷歌地图。

路段编号	道路名称	路段起点	路段方向	路段终点
101	东四环北路（主路）	机场高速	N→S	朝阳公园南路
102	东四环中路（主路）	朝阳公园南路	N→S	窑洼湖桥
103	东四环南路（主路）	窑洼湖桥	N→S	十八里店桥
104	东四环南路（主路）	十八里店桥	S→N	窑洼湖桥
105	东四环中路（主路）	窑洼湖桥	S→N	朝阳公园南路
106	东四环北路（主路）	朝阳公园南路	S→N	机场高速
201	东四环北路（内环辅路）	机场高速	N→S	朝阳公园路
202	东四环中路（内环辅路）	朝阳公园路	N→S	窑洼湖桥
203	东四环南路（内环辅路）	窑洼湖桥	N→S	十八里店桥
204	东四环南路（外环辅路）	十八里店桥	S→N	窑洼湖桥
205	东四环中路（外环辅路）	窑洼湖桥	S→N	姚家园新路
206	东四环北路（外环辅路）	姚家园新路	S→N	机场高速

表 3.1：北京市东四环双向主辅路路段编号与路段描述对应表

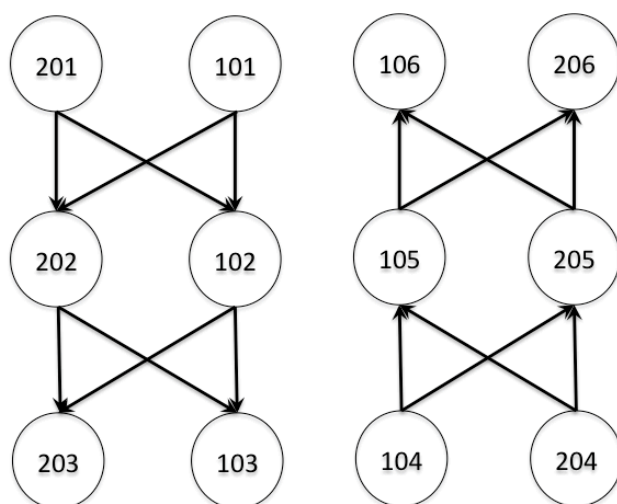


图 3.1：北京市东四环双向主辅路路段的路网结构图

数据中有少量缺失值，使用线性插值的方式补全。为使数值落在区间  $[0, 1]$  中，将所有速度值除以 100。使用前 10 天的数据作为训练集，第 11-12 天两天的数据作为测试集。将每相邻 12 组信号（每个信号为某时间间隔上所有 12 个路段的平均速度组成的向量，按照 4.4 的方法变形为  $12 \times 1 \times 1$  的张量）作为一个样本序列用于编码网络中历史数据的编码，将其后的连续 6 组信号作为预测网络的真实值序列，用于计算模型的损失函数和预测误差；这样，一共有 2863 个训练样本序列和 559 个测试样本序列。

### 3.2 超参数设置与实验细节

实验中先后使用了 GConvLSTM 网络层数分别为 1, 2, 3, 4, 5 的 5 种模型结构（模型各层通道数见表 3.2）。模型中的各参数使用标准差为 0.1 的零均值正态分布随机初始化，训练过程中 graph\_conv 运算的局部化参数设为  $K = 2$ 。对每一批训练样本，预测网络进行 6 个时间步的预测并与真实值序列比较，模型训练的目标为最小化序列 6 步预测的 L2 损失之和。优化方法采用

*RMSProp* (详见文献 [13]), 初始学习率为 0.01, 在每 1000 次迭代后下降 0.5 倍。模型一共迭代 3000 次, 每一次迭代的训练样本批量 (batch size) 为 25, 每 10 次迭代后计算一次模型对测试集序列做 6 步预测的相对误差率, 迭代完成后保留训练过程中测试集相对误差最小的测试集预测结果。

模型名称	网络各层通道数
GConvLSTM-1	12
GConvLSTM-2	$12 \times 12$
GConvLSTM-3	$12 \times 12 \times 12$
GConvLSTM-4	$6 \times 6 \times 12 \times 18$
GConvLSTM-5	$6 \times 6 \times 12 \times 12 \times 18$

表 3.2: 模型名称与对应的网络层数和各层通道数

在这里, 模型的超参数设置是经过实验得到的经验性结果。在下一小节中 (详见表 3.4), 将给出部分超参数试验的结果。

模型和实验使用 Python2.7 的 tensorflow 平台实现, 在使用一张 NVIDIA Tesla K80 计算卡的 GPU 服务器上运行。

为与 GConvLSTM 的预测表现做对照, 实验中还构建了 5 个基准模型: 线性回归 (记为 LR)、使用 Lasso 做变量选择的线性回归 (记为 Lasso)、支持向量机回归 (记为 SVR)、随机森林回归 (记为 RFR)、以及直接使用历史序列最后一个时间步的信号值作为预测 (记为 Copy-Last)。对于前四个基准模型, 设目标路段  $r$  ( $r = 1, 2, \dots, 12$ ) 上  $t+1$  时间步的速度与之前 1 小时 (12 个时间步) 的该路段的速度以及与该路段距离为 1 的所有邻居路段的速度相关, 即拟合模型

$$x_{t+1}^{(r)} = f(x_t(N(r)), x_{t-1}(N(r)), \dots, x_{t-11}(N(r))) + \epsilon,$$

其中  $x_t(N(r)) = (x_t^{(l)})_{l \in N(r)}$ ,  $N(r) = \{l : l = r \text{ 或 } (l, r) \in E \text{ 或 } (r, l) \in E\}$ ;  $\epsilon$  为随机误差。则由模型得到的  $x_{t+1}^{(r)}$  的预测值  $\hat{x}_{t+1}^{(r)}$  当做真实值带入拟合得到的模型  $\hat{f}$ , 可以迭代得到 6 步预测 (即 30 分钟预测)

$$\hat{x}_{t+6}^{(r)} = \hat{f}(\hat{x}_{t+5}(N(r)), \dots, \hat{x}_{t+1}(N(r)), x_t(N(r)), \dots, x_{t-6}(N(r))).$$

前四个基准模型使用 Python2.7 的 sklearn 模块提供的函数实现。另外, Lasso、SVR 和 RFR 的还有一些必要的参数设置, 详见表 3.3。

模型名称	参数设置与说明 <sup>4</sup>
Lasso	正则项系数 $\alpha = 1.0$
SVR	边界限度 $\epsilon = 0.1$ , 惩罚项系数 $C = 0.01$ ; 核函数使用线性函数
RFR	随机树数量 $n = 10$

表 3.3: Lasso, SVR, RFR 的参数设置

### 3.3 实验结果

对 5 种不同结构的 GConvLSTM 网络, 在 3000 次迭代后参数都达到收敛, 并且随着迭代次数的增加, 测试集上的误差呈现下降趋势, 每训练样本批次上的训练误差也大体呈下降趋势, 并且误差的波动幅度随学习率的阶梯式下降而递减。图 3.2 为 GConvLSTM-2 的训练误差与测试误差随迭代

次数变化的曲线, 其中水平的蓝色曲线为基准模型 Copy-Last 的平均测试误差  $MAPE_{Copy-Last} \approx 0.1312$ 。在第 2480 次迭代时预测网络达到最小平均测试误差  $MAPE_{Test} \approx 0.1129$ 。

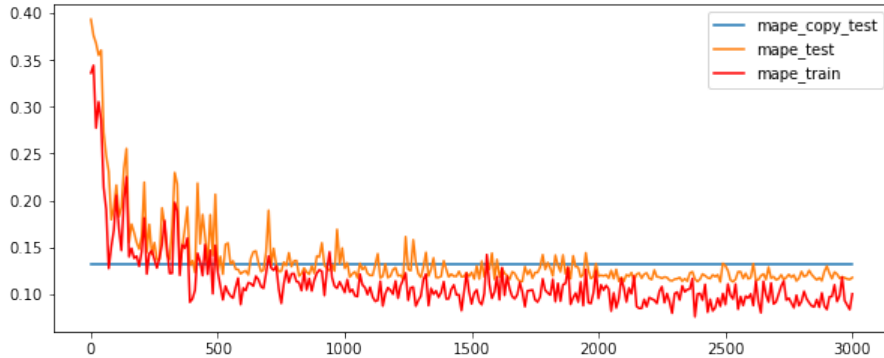


图 3.2: GConvLSTM-2 的误差-迭代次数曲线

事实上, 为了得到较好的模型表现, 实验过程中对模型的一些超参数——如图上卷积的局部化系数、迭代次数、初始学习率等——进行了调整并记录实验结果。以 GConvLSTM-1 为例, 表 3.4 中给出了对其部分超参数进行调整后模型测试集平均误差的对照试验结果。比较而言, 在本文中采用的模型超参数确实是效果较优的选择。

模型	$MAPE_{Test}$
GConvLSTM-1, 原设置	0.1149
GConvLSTM-1, 局部化系数 $K = 3$	0.1207
GConvLSTM-1, 2000 次迭代	0.1157
GConvLSTM-1, 4000 次迭代	0.1154
GConvLSTM-1, 初始学习率 0.02	0.1159
GConvLSTM-1, 初始学习率 0.03	0.1153

表 3.4: 对 GConvLSTM-1 的部分超参数设置的对照试验结果

表 3.5 记录了 5 个 GConvLSTM 网络和 5 个基准模型的训练集平均误差/最小训练批次平均误差和测试集平均误差。可以看到, 对训练集拟合效果最好的是随机森林回归 RFR, 而测试误差最小的是 GConvLSTM-2。综合来看, 5 个 GConvLSTM 网络和前 4 个基准模型的表现都超过了 Copy-Last, 而 GConvLSTM 网络的表现普遍好于基准模型, 其中, 表现最佳的 GConvLSTM-2 比基准模型的平均测试误差至少小 1% 左右, 也就是说 GConvLSTM 网络在预测准确性上比基准模型至少提高将近 1 个百分点、预测准确率接近 89%。

模型	$MAPE_{Train}(\min_i MAPE_{Batch_i})$	$MAPE_{Test}$
GConvLSTM-1	0.0709	0.1149
GConvLSTM-2	0.0757	0.1129
GConvLSTM-3	0.0738	0.1153
GConvLSTM-4	0.0778	0.1163
GConvLSTM-5	0.0717	0.1153
LR	0.0647	0.1217
Lasso	0.0655	0.1214
SVR	0.0628	0.1237
RFR	0.0268	0.1220
Copy-Last	—	0.1312

表 3.5: GConvLSTM网络和基准模型的训练误差与测试误差

尽管 GConvLSTM 在测试集误差方面相较于基准模型的提高不能算是十分显著，但是，通过对速度预测曲线和真实值曲线的绘图，可以发现，在对道路交通速度变化的预测上，GConvLSTM 比基准模型的响应要更加敏锐：前四个基准模型对速度值的突然升高和降低的预测具有比较明显的“时滞性”，在趋势上如同将真实速度曲线沿时间轴平移了 30 分钟，而 GConvLSTM 预测曲线的“时滞性”则不甚明显，尤其是在速度陡降的区间——这说明，至少在对道路拥堵的预测上，GConvLSTM 具有明显的优势。

图 3.3 为 Copy-Last, Lasso (基准模型中平均测试误差最小的模型) 和 GConvLSTM-2 在路段 101, 104, 204 上的预测值和真实值对比曲线，可以用于佐证 GConvLSTM 的上述优势。蓝色曲线表示真实值，橙色曲线表示预测值，从上至下分别为路段 101, 104, 204，从左至右分别为 Copy-Last, Lasso, GConvLSTM-2。

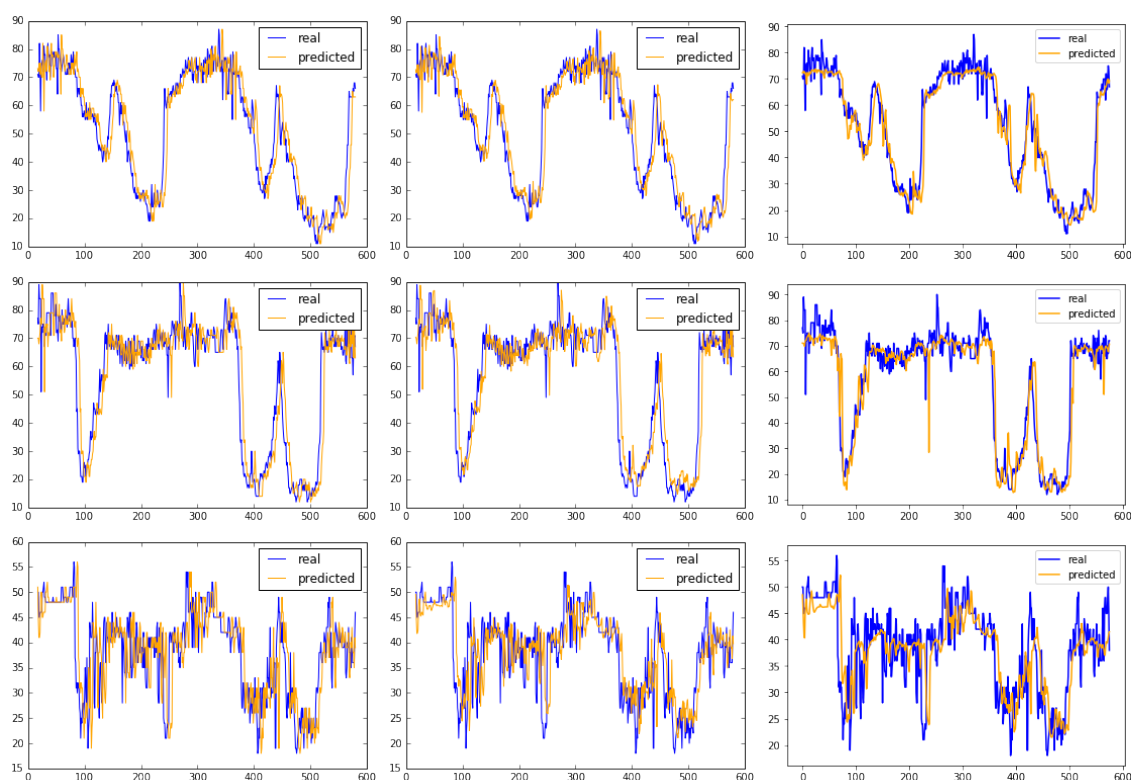


图 3.3: Copy-Last, Lasso, GConvLSTM-2 在路段 101, 104, 204 上的预测值和真实值对比

另外，考虑到数据测量和采集时的各种误差，预测曲线应当是在真实曲线的基础上去掉一部分随机误差后的相对平滑的曲线，这样得到的平均预测误差会更小，但是预测曲线也不能过于平滑，因为这样会失去对真实速度中的突变的敏锐性。也就是说，较为理想的预测曲线应介于过于平滑和过于“曲折”之间，而 GConvLSTM 得到的预测曲线则基本属于这种情况，并且，根据实验中观察得到的结果，其预测曲线的平滑程度与模型复杂度基本呈现负相关关系。图 3.4 以路段 105 为例，展示了层数分别为 2, 3, 4, 5 的 4 个模型（从左上至右下）。该路段在测试集第 1 天的夜晚到第 2 天的凌晨真实速度有很大的噪声水平，这主要是由于通行车辆太少、速度波动较大，而随着模型层数的增加，其预测曲线和真实数据的剧烈波动就越相似，但是这种相似性基本上是没有意义的；然而，对于测试集的其他时段的真实速度的突变和波动，随着模型层数的增加，其预测曲线与真实突变的相似性则是有意义的。

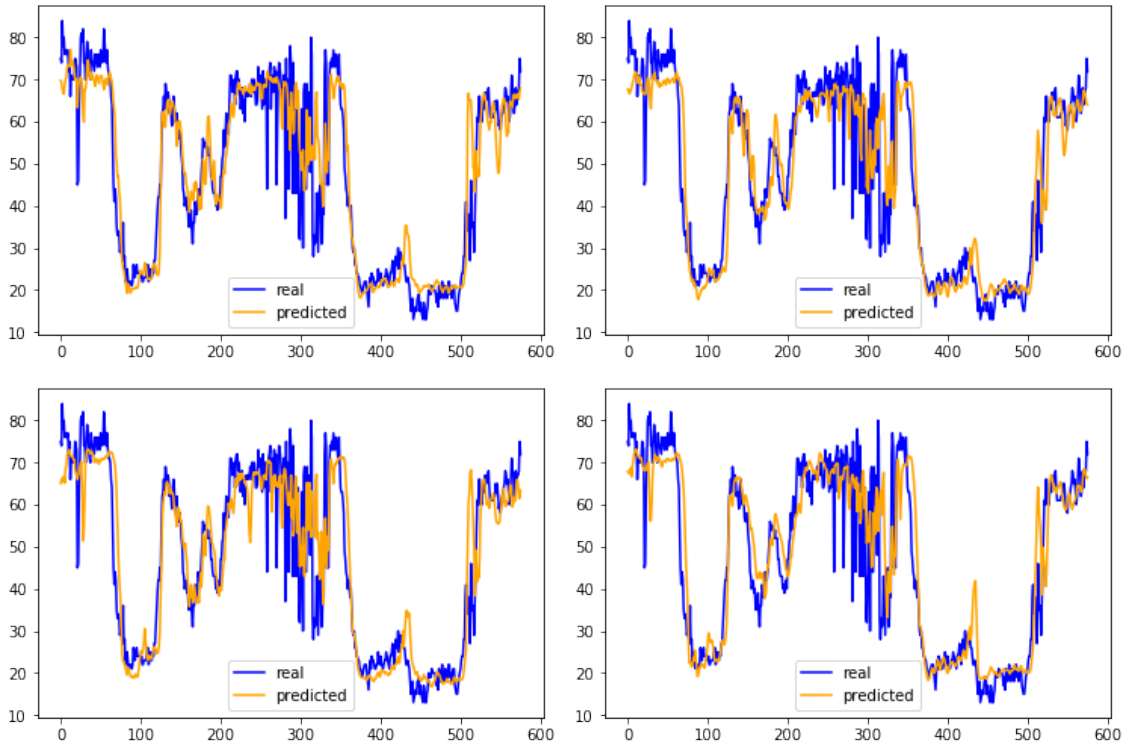


图 3.4: 层数为 2, 3, 4, 5 的 GConvLSTM 模型对路段 105 测试集速度的预测曲线

根据这样的特性，使用 GConvLSTM 对不同时段和不同性质的路段进行速度预测时，则可以灵活地选取模型的复杂度以得到更好的预测效果：在凌晨或深夜等因车流较少导致速度噪声较大的时段，就可以采用参数较少的 GConvLSTM 模型，以获得更为平均的速度估计；在车流量大或易发生拥堵的时段或路段，则可以考虑采用参数较多的 GConvLSTM 模型，以更好地捕捉速度突变。

## 4 结论与展望

本文将图上的卷积神经网络与卷积长短期记忆网络相结合，构建出图上的卷积长短期记忆网络 GConvLSTM，用于进行具有图结构的信号序列的多步预测，并将其运用于在北京市实际道路区域上的采集的实际速度数据集，获得了优于基准模型的半小时速度预测效果，并克服了基准模型预测速度的“时滞性”，并且预测值的噪声较小。此外，GConvLSTM 能够批量处理目标区域的全部路段速



度数据，并且一次性给出预测结果，无需针对所有单一路段分别建立模型，所以可以直接用于大面积区域甚至整个城市的所有道路的速度预测，这也是未来可以进行的工作之一。

由于时间和资源所限，本文作者对 GConvLSTM 的探索仍显不足，其预测准确率仍有提升的空间。在未来，可以从以下几个方向尝试继续改进 GConvLSTM 的预测表现：一，尝试新的网络结构，如增加层数、改变各层通道数等；二，尝试不同的损失函数、优化方法和学习策略；三，增加输入模型的信息（在数据可获得的条件下），如当日天气状况、道路长短周期性特征等。

另外，随着人工神经网络，特别是深度网络的迅猛发展，更加复杂和有效的信号序列预测模型层出不穷。未来也可以尝试使用文献 [14] 提出的混合循环神经网络（RNN Mixture）模型，对路网结构图中的顶点和边上的信号序列同时进行编码；还可以尝试文献 [15] 提出的使用对抗网络进行训练的卷积-长短期记忆-反卷积结构的网络，结合图上的卷积进行特征的捕捉并生成预测。



## 参考文献

- [1] Sun, Shiliang, Changshui Zhang, and Guoqiang Yu. A Bayesian network approach to traffic flow forecasting. *IEEE Transactions on intelligent transportation systems* 7.1 (2006): 124-132.
- [2] Min, Wanli, and Laura Wynter. Real-time road traffic prediction with spatio-temporal correlations. *Transportation Research Part C: Emerging Technologies* 19.4 (2011): 606-616.
- [3] Polson, Nicholas, and Vadim Sokolov. Deep learning predictors for traffic flows. *arXiv preprint arXiv:1604.04527* (2016).
- [4] Lv, Yisheng, et al. Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems* 16.2 (2015): 865-873.
- [5] Ma, Xiaolei, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies* 54 (2015): 187-197.
- [6] Jia, Yuhan, Jianping Wu, and Yiman Du. Traffic speed prediction using deep learning method. *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*. IEEE, 2016.
- [7] Xingjian, Shi, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems (NIPS)*. 2015.
- [8] Lotter, William, Gabriel Kreiman, and David Cox. Deep predictive coding networks for video prediction and unsupervised learning. *arXiv preprint arXiv:1605.08104* (2016).
- [9] Finn, Chelsea, Ian Goodfellow, and Sergey Levine. Unsupervised learning for physical interaction through video prediction. *Advances in Neural Information Processing Systems (NIPS)*. 2016.
- [10] Bruna, Joan, et al. Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv:1312.6203* (2013).
- [11] Niepert, Mathias, Mohamed Ahmed, and Konstantin Kutzkov. Learning convolutional neural networks for graphs. *Proceedings of the 33rd annual international conference on machine learning (ACM)*. 2016.
- [12] Defferrard, Michaël, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in Neural Information Processing Systems (NIPS)*. 2016.

- [13] Tieleman, Tijmen, and Geoffrey Hinton. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural networks for machine learning* 4.2 (2012).
- [14] Jain, Ashesh, et al. Structural-RNN: Deep learning on spatio-temporal graphs. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.
- [15] Lotter, William, Gabriel Kreiman, and David Cox. Unsupervised learning of visual structure using predictive generative networks. *arXiv preprint arXiv: 1511.06380* (2015).

## 致谢

感谢我的导师李铁军老师在本文的写作期间对我的指导和督促；作为交通大数据项目组的负责人，李铁军老师在两年前引导我进入了交通路况预测的研究领域，并一直在各方面给予我帮助和鼓励。感谢清华大学周培源应用数学研究中心的胡煜成老师，在数据预处理、路段分割和路网建立等重要工作上，他十分耐心地为我在提供宝贵的建议和反馈。感谢北京大学大数据研究院的朱占星老师，作为深度学习领域的专家，他指导我学习了人工神经网络与视频预测的基础知识，并在模型构建方面提出了深入和专业的见解。我还要感谢朱占星老师的硕士研究生殷浩腾同学，他在数据的收集、处理和探索中为我提供了很大的帮助，我在同他关于模型和实验方面的热烈讨论中也受益良多；另外，数学科学学院科学与工程计算系的博士研究生余冰同学无私地向我分享了神经网络构建的核心代码和使用 tensorflow 的技巧。最后，我要感谢数学科学学院和应用数学拔尖人才培养计划对我的培养与资金支持，感谢大数据研究院提供的数据和运算资源。

## 北京大学学位论文原创性声明和使用授权说明

### 原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名：                    日期：        年    月    日

### 学位论文使用授权说明

（必须装订在提交学校图书馆的印刷本）

本人完全了解北京大学关于收集、保存、使用学位论文的规定，即：

- 按照学校要求提交学位论文的印刷本和电子版本；
- 学校有权保存学位论文的印刷本和电子版，并提供目录检索与阅览服务，在校园网上提供服务；
- 学校可以采用影印、缩印、数字化或其它复制手段保存论文；
- 因某种特殊原因需要延迟发布学位论文电子版，授权学校☐一年/☐两年/☐三年以后，在校园网上全文发布。

（保密论文在解密后遵守此规定）

论文作者签名：                    导师签名：

日期：        年    月    日