

北京交通大学

硕士学位论文

基于深度学习的交通流量预测

Traffic Flow Prediction Based on Deep Learning

作者：魏中锐

导师：郭宇春

北京交通大学

2019 年 5 月

学位论文版权使用授权书

本学位论文作者完全了解北京交通大学有关保留、使用学位论文的规定。特授权北京交通大学可以将学位论文的全部或部分内容编入有关数据库进行检索，提供阅览服务，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。学校可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

（保密的学位论文在解密后适用本授权说明）

学位论文作者签名：



导师签名：



签字日期：2019年5月28日

签字日期：2019年5月28日

学校代码：10004

密级：公开

北京交通大学

硕士专业学位论文

基于深度学习的交通流量预测

Traffic Flow Prediction Based on Deep Learning

作者姓名：魏中锐

学 号：17125063

导师姓名：郭宇春

职 称：教授

工程硕士专业领域：电子与通信工程 学位级别：硕士

北京交通大学

2019 年 5 月

致谢

本论文的工作是在我的导师郭宇春教授的悉心指导下完成的，郭宇春教授科学的工作方法和严谨的治学态度对我产生了极大的影响和帮助。郭老师在工作中的认真负责的态度以及开朗豁达的人生理念都将对我今后的工作、学习和生活影响深远。在此衷心感谢两年来郭宇春老师对我的关心和指导。

同样感谢陈一帅老师对本次毕设的帮助和指导，在读研这两年内，多次和陈老师一起修改代码、讨论论文，每次交流都能产生新的想法。陈老师在工作 and 生活中饱满的热情以及负责的态度将会一直是我学习的榜样。

感谢实验室所有的老师，感谢赵永祥、李纯喜、郑宏云、张立军、孙强等老师在我研究生阶段对我的帮助，在此向各位老师表示衷心的感谢。同时感谢刘翔师兄、谭晓莹师姐、陈滨师兄、盛烨师姐、唐伟康师兄、胡伟师兄以及张大富师兄对我学习过程中的无私帮助。同样也感谢尹姜谊、于兹灏、张莹、周雪瑞、王亚珊、冯梦菲、艾方哲、苏健、宋云鹏、苏迪、杨晶晶、何美惠、曾显珣、陈泽龙、高志鹏、刘一健、胡安民等同学对我学习和生活中的帮助，感谢你们陪我度过这段难忘的学习生涯。还要诚挚地感谢国家自然科学基金（No. 61572071）的资助。

感谢一直无微不至地关心、支持我的父母和其他亲人朋友，正是他们热情的鼓励和默默的奉献，才使得我顺利地完成学业，成为社会的有用之才。

最后感谢北京交通大学这两年来对我的栽培，让我在研究生生涯中收获颇丰，感谢电子信息工程学院为我的学习提供了不断学习的机会，感谢网络与安全实验室为我提供了良好的环境去不断试验不断创新。

摘要

随着社会的发展,交通拥堵等交通问题已日渐显著。作为解决交通问题的有效手段之一,智能交通系统已经越来越受到人们的推崇。其中,交通流量预测是智能交通系统的基础,它通过利用历史数据对交通流量进行准确预测,从而对交通规划、用户出行提供准确的建议。然而,目前的交通流量预测的方法还存在很多的不足。大部分的交通流量预测方法仅基于流量的时间序列特征信息进行预测,这样会造成信息的缺失从而降低预测的准确性;其它的预测方法虽综合利用了特定地点的时空特征信息,但空间特征提取方法并不完善,提取出的空间特征信息的准确度有待提升,因此会对预测的准确性带来不良影响。

针对上述问题,本文提出了一种基于深度学习的交通流量预测方法,通过对数据的处理与整形,利用图卷积网络与长短期记忆网络提取交通流量的时空特征,对其进行预测,并结合时间周期、天气信息等其他相关因素特征对交通流量进行预测的结果,实现更准确的预测。本文的主要贡献如下:

(1) 针对传统的交通流量预测空间特征提取方法的不足,提出了一种更实用、更符合实际交通网络结构的时空特征提取方法。该方法的特色在于:它将路网视为拓扑图,其中道路上的交通流量作为边的特征,利用图论中线图转换的方法,将路网拓扑转换为路邻接拓扑,道路上的交通流量变为节点信息特征,然后将路邻接拓扑矩阵与节点特征矩阵送入图卷积网络中,实现对交通流量空间特征的合理提取与利用。在北京市出租车数据集上的实验评估结果显示,该方法相对于 CNN+LSTM 准确率有显著的提升,均方误差为 1.8774,预测误差降低了 56.9%。

(2) 通过结合深度学习提取的时空特征的预测结果与其他相关特征的预测结果,设计了一种基于深度学习的交通流量预测系统。该系统通过合理提取交通流量的时空特征进行预测,然后将该结果与利用交通流量有关的其他特征(周期特征、天气特征)的预测结果结合,得到最终的交通流量预测结果。在北京市出租车数据集上的实验评估结果显示,本文提出的交通流量预测系统相比于以往的方法准确率有了明显提升,均方误差为 1.6203,相比于 CNN+LSTM 预测误差降低了 62.8%,相比于(1)中提出的预测方法预测误差降低了 13.7%。

本文提出的流量预测方法具有通用性,能够用于道路网络和信息网络的时空序列预测问题,因此有一定的理论研究价值。本文提出的流量预测方法和系统能够准确预测道路交通流量,有着很强的现实意义和实际价值。

图 30 幅,表 7 个,参考文献 42 篇。

关键词: 深度学习; 交通流量预测; 线图; 图卷积网络

ABSTRACT

With the development of society, traffic problems such as traffic congestion have become increasingly prominent. As one of the effective means to solve traffic problems, intelligent transportation systems have become more and more popular. Noteworthy, traffic flow forecasting is the basis of intelligent transportation system. It makes accurate predictions on traffic flow using historical data to provide appropriate advice for traffic planning and passenger travel. However, there are still many shortcomings in current methods traffic flow prediction.

Most of the traffic flow prediction methods are mainly based on the time series characteristics, which will result in the lack of information and reduce the accuracy of the prediction. Other prediction methods comprehensively utilize the spatio-temporal feature of a specific location, but the spatial feature extraction method is not perfect, which means the accuracy of the extracted information needs to be improved, and this insufficiency leads to the low prediction ability of the method.

Aiming at solving the forementioned problems, this thesis proposes a traffic flow prediction method based on deep learning. After the processing and shaping of data, the graph convolutional network and the long short-term memory network are used to extract the spatial and temporal characteristics of traffic flow, respectively, and then an initial prediction result is obtained, integrating which with two versions of predicting results using time-period and weather information, a final result which is more accurate is achieved. The main contributions of this thesis are as follows:

(1) To get over the shortcomings of traditional spatial feature extraction methods for traditional traffic flow prediction, a spatial feature extraction method which is more practical is proposed and it manifests better conformability to the actual traffic network structure. The method is characterized in that it regards the road network as a topological map, in which the traffic flow on the road is adopted as the feature of the edge, whereas the obtained feature turns to the node feature with the application of line-graph theory illustrated in graph theory, through which the road network topology can be transformed into a road adjacency topology, and then the road adjacency topology matrix and the node feature matrix are sent into the graph convolution network to realize a reasonable extraction and utilization of the traffic flow spatial characteristics. The experimental evaluation results based on the taxi

vehicle dataset in Beijing show that using the proposed method, the traffic flow prediction accuracy is significantly improved compared with the method based on CNN+LSTM, and the prediction error is reduced by 56.9%.

(2) Combining the spatio-temporal features extracted with deep learning method with other related features, a novel traffic flow prediction system based on deep learning is proposed. The system predicts the traffic by reasonably extracting the spatio-temporal characteristics of traffic flow, and then integrates the prediction results with other characteristics related to traffic flow (periodic characteristics, weather characteristics) to obtain the final traffic flow prediction results. The experimental evaluation results based on the taxi vehicle dataset in Beijing show that the traffic flow prediction system proposed in this thesis achieves a significantly-improved predicting accuracy compared with previous method. For instance, compared with the prediction method based on CNN+LSTM, the prediction error in our system is reduced by 62.8 %, and with using the method illustrated in (1), the prediction error of our system is reduced by 13.7%.

The flow prediction method proposed in this thesis is universal and can be applied in similar spatio-temporal sequence prediction problems on traffic network and information network, so it has certain theoretical research value; in addition, the traffic prediction method and system proposed in this thesis can accurately predict road traffic flow, which means a lot from the perspective of reality and practical application.

KEYWORDS : Deep learning; Traffic flow prediction; Line graph; Graph convolution network

目录

摘要.....	III
ABSTRACT	IV
1 引言	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	3
1.2.1 传统交通流量预测.....	3
1.2.2 基于机器学习的交通流量预测.....	4
1.2.3 基于深度学习的交通流量预测.....	4
1.3 主要工作及研究内容与贡献.....	6
1.4 本论文的组织结构.....	8
2 论文相关知识介绍	10
2.1 开发工具介绍.....	10
2.1.1 Python 语言	10
2.1.2 Numpy 数组运算库.....	11
2.1.3 Pandas 数据处理库	11
2.1.4 Statsmodels 统计模型库	12
2.1.5 Hmmlern 机器学习库	12
2.1.6 TensorFlow 深度学习框架	13
2.2 传统时间序列预测方法.....	14
2.2.1 自回归模型.....	14
2.2.2 向量自回归模型.....	14
2.2.3 移动平均模型.....	15
2.2.4 自回归滑动平均模型.....	15
2.2.5 差分整合移动平均自回归模型.....	15
2.3 机器学习算法.....	17
2.3.1 隐马尔可夫模型.....	17
2.4 深度学习算法.....	18
2.4.1 循环神经网络.....	19
2.4.2 卷积神经网络.....	22
2.4.3 图卷积网络.....	23
2.5 线图.....	24
2.6 本章小结.....	25
3 系统设计与数据分析处理	26
3.1 系统设计.....	26
3.2 数据分析.....	27
3.3 数据预处理.....	28

3.4 本章小结.....	31
4 基于深度学习的交通流量预测.....	32
4.1 交通流量空间特征提取及预测.....	32
4.1.1 线图转换.....	33
4.1.2 空间特征提取及预测.....	34
4.2 交通流量时间特征提取及预测.....	35
4.3 交通流量其他特征提取及预测.....	35
4.3.1 周期特征提取及预测.....	36
4.3.2 其他特征提取及预测.....	37
4.4 预测结果整合.....	37
4.5 本章小结.....	37
5 实验对比与结果分析.....	38
5.1 参数设置与结果展示.....	38
5.1.1 系统参数.....	38
5.1.2 实验结果展示.....	39
5.2 模型结果对比与分析.....	40
5.2.1 交通流量预测基线模型.....	40
5.2.2 结果对比分析.....	47
5.3 各模块结果与功能分析.....	49
5.3.1 空间特征提取模块.....	49
5.3.2 其他特征提取模块.....	49
5.4 本章小结.....	51
6 总结及展望.....	52
6.1 论文工作总结.....	52
6.2 未来工作展望.....	53
参考文献.....	55
作者简历及攻读硕士/博士学位期间取得的研究成果.....	58
独创性声明.....	59
学位论文数据集.....	60

1 引言

1.1 研究背景及意义

随着不断攀升的机动车保有量，道路交通系统所面临的形势也越来越严峻。经济的增长、急剧的工业化与城市化、私人汽车的普及，这些都导致了越来越多的交通拥堵甚至于是交通事故。因此，交通问题是随着社会发展人类必须要解决的一个重要问题。在国务院提出的《国家中长期科学和技术发展规划纲要(2006—2020年)》^[1]中也同样指明我们要实现交通信息共享和各种交通方式的有效衔接，重点开发综合交通运输信息平台和信息资源共享技术。

智能交通系统(Intelligent Traffic System, ITS)^{[2][3][4]}是一种先进的交通管理控制系统,它旨在结合数字通信与计算机网络等先进的信息技术来对交通系统提供与原始的运输方式和交通管理不同的创新型服务,这能够使得用户可以更好地了解交通的情况,并且能够使用更安全、更协调、更智能的交通运输网络。在智能交通系统中,交通流量预测占有显著的位置,交通流量预测是智能交通系统实现的基础和关键所在,科学的管理方法必须有数据的支持,只有在历史交通数据上进行高精度的交通流量预测,才能向出行者提供更为有效的道路交通信息,为动态交通规划提供数据支撑,从而达到缓解道路交通拥堵和方便人们出行的目的。因此,交通流量预测就显得十分必要,预测结果可以指导交管部门对道路流量进行科学的管理,从而提高道路资源的利用率。

对于道路上的交通流量预测这一重要问题,现阶段主要根据获取数据类型不同分为两大部分,其中占主要部分的是根据对道路上的路网摄像头所采集的道路上车辆的行驶速度进行预测进而对道路上的交通流量进行反应,其余部分是根据对道路上的路网摄像头中采集的图像中车辆进行识别从而统计道路上的交通流量或直接使用车辆信息数据集产生道路上的交通流量。目前在交通流量预测方面并没有一个完整的体系,关于交通流量预测的研究仍有许多难题需要克服。一方面交通流量数据巨大、信息丰富,采集到的道路交通信息包含很多的信息,往往所要需要的内容只占全部道路交通信息的很少一部分,如果对采集到的所有的道路交通信息直接进行处理,不但会增加计算成本,而且容易导致错误率增加,所以需要对采集到的道路交通信息进行处理、整形。另一方面要对道路上的交通流量进行预测,不仅要处理、整形得到合适交通流量数据,还要选取合适的模型对交通流量进行预测,选取的模型不仅仅决定了所需要的道路交通流量数据的形

式,还决定了最终交通流量预测结果的准确程度,这也增加了交通流量预测的难度。目前关于交通流量预测的研究工作大多是只基于单一地点的交通流量数据序列来进行预测^{[10][11][12][27][28][33]},即将时间进行均匀切分,统计各个时间切片中的车辆流量,从而得到一个特定位置的交通流量时间序列,以此来对该地点的交通流量进行预测。然而这类方法只能对所选取的特定位置的交通流量进行预测,只利用了所在位置的时间信息,并未利用该位置的空间信息,这样预测的结果很容易由于信息缺失造成预测准确性的偏移。有部分研究同时利用了道路上交通流量的时空特征来对道路上的交通流量进行预测^{[33][35]},但是对空间特征的提取力度过于粗糙,同样会对道路上的交通流量预测结果的准确性有不良影响。图 1-1 展示了道路上车辆交通流量分布热力图,可以看出道路上的交通流量与空间位置有很大的关系,因此合理的提取道路中的空间流量信息就显得十分必要。因此,共同利用道路上交通流量数据的时间特征与空间特征来预测道路上的交通流量是很必要的,而合理的提取道路交通流量的空间特征是准确预测道路上的交通流量的基础。

因此,通过对实际数据特征的分析,本文设计了一种基于深度学习的交通流量预测系统,旨在通过利用从交通流量信息中提取的空间特征以及时间特征结合周期特征和其他特征共同作用来实现交通流量的预测。与现有的方法不同,本文不仅仅整合了交通流量数据的空间特征与时间特征,同时还结合其他信息,共同作用来对交通流量进行预测。同时,在提取道路上交通流量的空间特征时,巧妙的利用线图^[5](line-graph)的方法对数据进行转换,并选取了一种更为合理的提取方法,大大提高了对道路上交通流量预测的准确性。因此,本文所设计的交通流量预测系统考虑的方向全面,并巧妙的解决了难以合理的提取道路上交通流量的空间特征这一问题,能够更好的对道路上的交通流量进行预测。

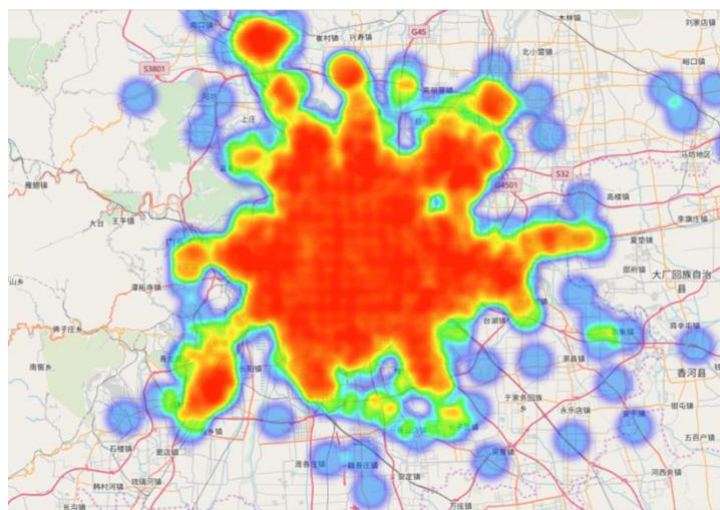


图 1-1 交通流量分布热力图
Figure 1-1 Distribution Heat Map of Traffic Flow

1.2 国内外研究现状

随着社会的发展与进步,智能交通系统已经在显示生活中有了很广泛的应用,其在交通流量预测、交通系统优化、车辆优先控制、交通系统安全等方面^{[6][7]}都发挥着重要且独特的作用。其中,对于交通流量预测问题,国内外学者对其进行了大量研究,目前关于交通流量预测问题的研究主要集中在特征提取与模型选择这两个问题上。特征提取主要是针对既有的交通流量数据,提取其中所含有的信息,主要是时间特征信息与空间特征信息两方面,这里的时间特征信息提取的主要是是相邻时间间隔的数据所隐藏的信息,还可以提取周期性的时间特征信息,而空间特征信息提取的主要是道路的空间结构特征信息。而模型选择方面主要是针对特征提取的种类与方式的不同来选取合适的模型来对道路上的交通流量进行预测。本节将会介绍交通流量预测研究的现状,包含传统交通流量预测方法、基于机器学习的交通流量预测方法、基于深度学习的交通流量预测方法等,并分析这些以往交通流量预测研究的特点。

1.2.1 传统交通流量预测

传统的交通流量预测方式主要是将交通流量数据整形成一个交通流量的时间序列,即将交通流量预测问题划归为时间序列预测问题。时间序列预测主要是根据已有的时间序列数据去预测该数据未来的变化趋势,通过对已有的时间序列进行分析,提取其中的变化特征,并将其推广到要预测的数据上,以此来对未来数据的变化进行预测。

在以往的研究过程中,文献[8]提出了一种基于自回归滑动平均(Auto-Regressive and Moving Average, ARMA)模型的城市轨道交通客流量预测模型,该模型利用 ARMA 模型对天津地铁客流量进行拟合与预测。文献[9]提出了一种利用 K 近邻(K-NN)方法作为预测算法来回避参数预测方法中固有的一些问题并将其与简单的单变量线性时间序列预测进行比较,结果发现 K-NN 方法与线性时间序列方法的表现相当,而更大的数据库可以提高 K-NN 方法的准确性。文献[10]提出了一种基于季节的自回归移动平均模型(Seasonal Auto-Regressive Integrated Moving Average, SARIMA),该模型相比传统的 ARIMA 考虑了不同季节对车流量的影响,仅使用少量的输入数据预测短期流量, SARIMA 模型克服了 ARIMA 模型在交通流量预测中的适用性问题。可是传统的交通流量预测方法大都需要较大的数据量且数据需要平稳,因此对原始数据要求较高。

1.2.2 基于机器学习的交通流量预测

依托于不断发展的大数据，机器学习也日趋壮大。大数据是人工智能的基础，而应用大数据来实现人工智能的过程就需要机器学习来发挥作用，机器学习是使得人工智能得以存在的基石，是人工智能中最重要的一环。在交通流量预测方面机器学习也占有不可或缺的一席位置，机器学习的加入，使得利用更少量的数据来对道路上的交通流量进行更为准确的预测成为了可能。

在以往的交通流量预测方法中，机器学习算法也是占有重要地位的。文献[11]提出了一种基于支持向量回归机的交通状态短时预测方法，利用训练完成的支持向量回归机，来对下一时段道路上的交通流量进行预测，并与真实值进行对比来评估模型效果的优劣，结果说明了模型的可行性。文献[12]提出了一种利用隐马尔可夫模型（Hidden Markov Model, HMM）来对高峰期的短期高速公路交通流量进行预测，实验结果表明 HMM 方法成功地实现了对高峰期的短期交通状况进行了预测，并考虑了交通状况的固有随机性。而该隐马尔可夫模型中使用的是一阶马尔可夫，即只考虑上一状态的影响，这会丢失很多时间信息。

1.2.3 基于深度学习的交通流量预测

在机器学习的范围内，有着一类更为细致且重要的一部分—深度学习。深度学习相对机器学习来说，它除了可以学习特征和任务之间的关联以外，还能自动从简单特征中提取更加复杂的特征。深度学习可以更有效的进行特征提取进而处理一些更为复杂的问题，在网络加深的过程中，通过对底层提取的浅特征重新提取整合，从而自动生成复杂的深层特征，并利用这些组合特征解决问题。

当下，深度神经网络(Deep Neural Networks, DNNs)发展的如火如荼，在计算机视觉^{[13][14]}、语音识别^{[20][22]}、自然语言处理^{[23][26]}等人工智能相关领域都颇有建树。受到深度学习在多个领域取得成功的启发，很多研究人员也开始尝试将深度学习的方法用于解决车流量预测问题。现今能够对道路上的交通流量进行一个好的预测的算法大都是基于深度学习与深度神经网络的算法，因此相比较于传统的数据挖掘方法，研究人员现如今更加青睐于基于深度学习的算法。

文献[27]提出了一种基于深度神经网络（DNN）的交通流量预测模型（DNN-BTF）来提高预测精度，该模型充分利用了道路上交通流量的以每周和每日为时间周期的特征和时空特征，并利用开放式数据库对算法性能进行验证，结果显示该方法优于最先进的算法。文献[28]提出了一种设计了一种基于深度学习的交通流量预测方法，该方法融合了卷积神经网络和长短时记忆神经网络两种

网络结构,卷积神经网络用于提取空间特征分量,长短时记忆神经网络综合提取出来的空间特征分量做序列预测。现阶段方法大都注重于基于时间尺度上的流量预测,而忽视空间尺度上特征,本文虽然利用卷积神经网络提取道路上交通流量的空间特征,但是路网与图片有很大区别,卷积神经网络在提取空间特征时会引入很多不必要特征,所加入的特征有可能与实际道路情况不符,进而影响最终的预测结果。文献[29]提出了一种新颖的混合卷积长短期记忆神经网络(ConvLSTM)来预测关键路段交通流量的变化。通过时空相关算法来识别对子网具有最强影响的关键路段,以次来预测整个网络的未来流量状态。实验结果表明,CRS-ConvLSTM 优于主流深度学习(DL)方法,适用于考虑关键路段的情况,并具有普适性。文献[30]对深度残差网络进行了改进并构建了一个以前的研究人员几乎不使用的动态模型,实验结果表明该模型不仅可以有效地训练,而且还具有更高的准确性并更适合实际应用。文献[31]提出了一种预测人类移动的模型—DeepMove,该模型通过多模态嵌入循环神经网络对复杂的人类移动性的多个因素进行顺序捕捉转变,并设计两种机制的 attention 模块来捕获多层次的周期性,最终在多个真实移动数据集上进行实验,结果表明 DeepMove 比最先进的模型的性能要高出 10%以上。文献[32]提出一种基于长短期记忆网络(Long Short-Term Memory, LSTM)的可以结合多人信息来预测场景中的人类轨迹的模型,为每条轨迹使用一个 LSTM,并通过引入新的社交池层来共享 LSTM 之间的信息,成功预测了由社交互动引起的各种非线性行为,并在两个公开可用的数据集上进行试验,预测效果优于最先进的方法。文献[33]提出了一种新的模型--LC-RNN,它通过合理地整合它们来利用 RNN 和 CNN 模型,以便学习更有意义的时间序列模式,以适应周围区域的交通动态,进而对道路上的速度进行预测。并提出了一种网络嵌入式卷积结构来捕获拓扑感知特征,最终与其他信息的融合,包括周期性和背景因素,进一步提高对道路上速度预测准确性。并在两个真实数据集上进行试验,结果显示 LC-RNN 优于七种现有的算法。文献[34]提出了一种深层叠双向单向 LSTM(SBU-LSTM)神经网络,该神经网络考虑时间序列数据的前向和后向依赖性,预测整个网络的流量速度,其中利用双向 LSTM(BDLSM)层从历史数据中捕获空间特征和双向时间依赖性。与经典模型和最先进模型的比较,结果显示 SBU-LSTM 神经网络在准确性和鲁棒性方面都体现出了卓越的性能。文献[35]提出了一种时空循环卷积网络(SRCN),该网络可以基于先前的场景预测对象的未来运动,所提出的时空循环卷积网络(SRCN)继承了深度卷积神经网络(DCNN)和长期短期记忆(LSTM)神经网络的优点,可以捕获网络范围流量的空间依赖性并学习时间动态,最终对道路上的今天流量进行预测,实验结果表明 SRCN 在短期和长期交通预测方面都优于其他基于深度学习的算法。文

文献[36]提出了一种基于深度学习的预测算法 DeepTFP, 以预测城市每条交通道路上的交通流量, 通过使用三个深度残差神经网络对交通流量的时间紧密度、周期和趋势特性进行建模, 最后在英格兰交通系统的移动时间序列数据集上进行实验, 结果显示 DeepTFP 在预测精度上优于基于长期短期记忆(LSTM)体系结构的方法。文献[37]提出了一种基于深度学习的方法--ST-ResNet, 用于集体预测城市每个区域的人群流入和流出, 使用残差神经网络来模拟人群流量的时间紧密度、周期和趋势属性, 并对每个属性设计了一个残余卷积单元, 最后基于数据动态地聚合三个残余神经网络的输出, 并在北京和纽约的两个数据集上进行实验, 实验结果表明所提出的 ST-ResNet 优于六种众所周知的方法。文献[38]提出了一种新的深度学习框架--时空图卷积网络 (STGCN), 设计一种包含图卷积网络与门控卷积神经网络的时空卷积模块, 对道路上流量速度进行预测, 在两个真实数据集上进行实验, 实验结果超越了其他最先进的方法, 并拥有更快的训练速度, 更好的融合性, 以及更少的参数、灵活性和可扩展性等优点。文献[39]提出了一种基于长短期记忆 (LSTM) 网络的新型交通预测模型。与传统的预测模型不同, 所提出的 LSTM 网络通过由许多存储器单元组成的二维网络来考虑交通系统中的时空相关性。与其他代表性预测模型比较验证了所提出的新型交通预测模型可以实现更好的性能。文献[40]提出了一种 Hetero-ConvLSTM 框架, 这个框架在基本的 ConvLSTM 模型之上实现了一些新颖的想法: 合并空间图特征和空间模型集合, 解决数据中的空间异质性挑战。在实验中得出所提出的框架可以进行相当准确的预测, 并有着明显高于基线方法的预测准确性。以上的方法随有一个不错的效果, 但大都体现在对速度的预测上, 而无法应用到道路上的交通流量预测上。与此同时, 很多方法在空间特征的提取上还不是很符合实际, 并且还未考虑其他特征对流量预测的影响。

1.3 主要工作及研究内容与贡献

本文主要研究基于深度学习的交通流量预测问题, 为了将问题具像化, 本文利用现有的车辆信息数据集作为研究对象进行实验, 对道路上的交通流量进行预测。

本文的主要工作及研究内容如下:

(1) 创立实验环境, 选取所用框架。在 Linux 系统上搭建了本次的实验平台, 安装了深度学习框架 TensorFlow。本文主要基于 TensorFlow 框架, 了解、学习并使用 TensorFlow 进行开发, 阅读 TensorFlow 文档和源码, 编程实现基于 TensorFlow 的交通流量预测深度学习算法。

(2) 分析实验数据，实现数据预处理。深入分析所拥有的车辆数据集，了解车辆数据集的统计信息，熟悉车辆数据集情况，并对车辆数据集实施清洗、挖掘等操作删去无用值、补全或删除去缺失值，整理成可用车辆数据集。

(3) 时空特征提取及预测。本文将特征提取及预测分为三部分。空间特征提取是本文基于深度学习的交通流量预测方法的第一步流程，针对具体的问题，本文利用了一种新的空间特征转换与提取方法，即利用线图（line-graph）实现节点连边转换，得到路网结构拓扑邻接矩阵与将道路流量特征进行映射得到的节点特征，再利用图卷积网络进行空间特征提取。下一步是时间特征提取及预测，将已经提取完的空间特征的数据进行整形，得到已经包含空间特征的时间序列数据，并将其送入循环神经网络中的一类--长短期记忆网络（LSTM, Long Short-Term Memory）来提取时间特征，进而预测交通流量。

(4) 其他特征提取及预测。分别提取交通流量数据的时间周期特征与其他特征（如天气信息），并分别以提取的两类其他特征来对道路上的交通流量进行预测。

(5) 多维预测结果整合。通过对通过对时空特征提取后预测的结果与其他特征提取后预测的结果进行整合，进而得到所设计的交通流量预测系统的最终预测结果。

(6) 实验结果对比分析。在真实数据上对本文所提出的交通流量预测系统进行效果评估，并选取以往的有代表性的以及有可比性的交通流量预测方法在相同数据集上进行试验并进行结果对比，分析本文提出的系统的优点。并对本文所提出的交通流量预测系统中的各个主要模块进行评估，分析各模块性能。

在实现本次交通流量预测系统实现的过程中，本文做出了以下贡献：

(1) 提出了一种更切合实际的空间特征提取方法。本文利用现有车辆信息数据集，将其经过大量的数据清洗及整形变换为道路上交通流量数据。根据数据的结构特点以及本文所选取模型的特点，将路网视为拓扑图，利用线图（line-graph）的方式进行节点连边转换，此时道路上的车辆流量特征即转换为节点信息特征，得到节点特征矩阵，并生成转换后拓扑图的邻接矩阵，此邻接矩阵即表示路网中道路的邻接关系。将得到的节点特征矩阵与道路邻接关系邻接矩阵一起输入图卷积网络，从而提取道路上交通流量的空间特征。本文这样做的优点在于将现实中的路网看作拓扑图，更符合实际的道路空间特征；同时利用线图的方式进行转换，将道路上的特征值转换为节点特征信息，便于输入图卷积网络进行运算。在真实数据的实验中，该方法对道路上交通流量进行预测的预测结果的均方误差（RMSE）达到 1.8774。

(2) 设计并实现了基于深度学习的交通流量预测系统。以往的交通流量预测

方法大多只使用了交通流量信息的时间特征或者粗略的提取交通流量信息的空间特征并将其与交通流量信息的时间特征进行融合再对交通流量进行预测,而本文则是更合理、更切合实际的提取了交通流量信息的空间特征并将其与交通流量信息的时间特征进行融合进行预测再结合交通流量数据的时间周期信息与一些其他信息的预测结果,作为最终的交通流量系统预测结果,以此来提高交通流量预测的准确性。在真实数据的实验中,本文设计的基于深度学习的交通流量预测系统的预测结果均方误差(RMSE)达到 1.6203。

1.4 本论文的组织结构

本文的组织结构如下:

第二章详细介绍了本文实验中所使用的开发工具的安装及使用方法,包括相应的数据处理工具库、机器学习工具库以及深度学习平台。同时对相关的背景知识进行了梳理,从传统的基于统计学的时间序列预测方法到与时间序列交通流量预测相关的机器学习算法与深度学习算法,并简要的介绍了机器学习与深度学习的相关知识。之后介绍了本文所使用的数学原理,最后对本章的内容进行了总结归纳。

第三章对本文所设计的系统进行了剖析并对实验数据进行分析并进行预处理。首先阐述本文所设计的系统的目的以及该系统所关注的问题,根据以上分析对整个系统的流程进行设计,并阐明了系统设计过程中各个组成模块以及各个模块的作用。接下来针对所拥有的实验数据,首先介绍了数据集的内容及结构,并对数据集进行统计处理,查看数据集的统计学规律,以验证提出的问题的准确性,重点介绍了对数据集进行预处理的方法与工作步骤。最后对本章进行了总结。

第四章介绍了本文提出的基于深度学习的交通流量预测系统。主要介绍了其中所包含的提取道路上的交通流量信息的空间特征、时间特征与其他特征的方法与过程。首先对本章的工作进行整理分析,即提取整形后的交通流量数据中的空间特征、时间特征与其他特征然后进行预测,并对特征提取顺序进行排序。之后解释了提取特征时选择图卷积网络的原因并介绍了使用图卷积网络(GCN)提取交通流量数据中空间特征的方法,利用线图(line-graph)方法进行转换,并生成道路邻接拓扑矩阵与节点特征矩阵,将转换后的结果作为图卷积网络(GCN)的输入,用以提取道路上的交通流量的空间特征。这之后介绍使用长短期记忆网络提取交通流量数据中时间特征的过程。接着介绍了提取时间周期特征与其他特征的意义与方法。最后对本章的内容进行了总结。

第五章主要对本文所设计的交通流量预测模型进行试验评估,利用选取的评

价指标,对本文所设计的交通流量预测模型的效果进行评价。同时介绍了以往的交通流量预测方法在本文数据集上的实现,并将这些方法的效果利用同样的评价指标进行评价,并将本文所设计的交通流量预测模型的效果与以往的交通流量预测方法的效果进行比较,分析产生结果的原因并结合模型算法的本质加以解释。同时,对本文所设计的交通流量预测系统进行剖析,分析本文所设计的交通流量预测系统中各部分的功能以及其对交通流量预测系统的影响,并加以解释。最后对本章内容进行总结概括。

第六章主要对整篇论文进行了全面的概括总结并对下一步的研究方向进行了展望。主要展示了本文所解决的挑战与问题并列举了本文所做的贡献,接下来对本文的下一步研究方向进行分析展望。

2 论文相关知识介绍

本章将会介绍本文实验中所使用的开发工具及所涉及的相关知识，以便于读者对本文所做的工作进行更为细致、全面的了解。首先介绍了本文实验中所使用的开发工具，并阐述了它们的安装及使用方法，其中包括相应的数据处理工具库、机器学习工具库以及深度学习平台。接下来对相关的背景知识进行介绍，首先介绍了传统的时间序列预测方法，如 VAR、ARIMA 等，然后对机器学习进行了大致说明并介绍了一种常见的常用于时间序列预测的机器学习算法——隐马尔可夫模型（HMM 模型），接下来简要描述了深度学习的发展并介绍了深度学习中的循环神经网络（RNN）、卷积神经网络（CNN）和图卷积网络（GCN）。这之后介绍了图论中的线图（line-graph）的概念，这将用于接下来的数据处理工作中。在本章最后对本章内容进行小结。

2.1 开发工具介绍

2.1.1 Python 语言

Python 是一种解释性高的，高级的通用编程语言。它由 Guido van Rossum 创建并于 1991 年首次发布，其设计理念强调代码的可读性，特别是使用空白的重要性。它提供的结构可以实现小规模和大规模的清晰编程。与 Java 或 C++ 相比，Python 中的语法有助于程序员以更少的步骤进行编码。

近些年来 Python 在其他编程语言（如 C，C++ 和 Java）中排名第一，并且被程序员广泛使用。产生这种结果的原因是 Python 语言在软件开发中具有多样化的应用，例如游戏，Web 框架和应用程序，语言开发，原型设计，图形设计应用程序等。这使得该语言比业内使用的其他编程语言更多。其优点包括广泛的支持库、强大的集成功能、更好的提高程序员的工作效率、强大的生产率等。

在安装 Python 时，一般都先在电脑上安装 Anaconda，因为 Anaconda 让使用者省时省心，并且 Anaconda 是一个分析利器。Anaconda 通过管理工具包、开发环境、Python 版本，大大简化了你的工作流程。不仅可以方便地安装、更新、卸载工具包，而且安装时能自动安装相应的依赖包，同时还能使用不同的虚拟环境隔离不同要求的项目。在 Anaconda 官网中是这么宣传自己的：适用于企业级大数据分析的 Python 工具。它不仅可以做数据分析，甚至可以用在大数据和人

工智能领域。同时 Anaconda 中可以共存同时管理两个版本的 Python。

本文的开发内容，包括数据处理与深度学习算法两方面，都是使用 Python 完成的，只需要调取不同的工具包以实现不同的功能。

2.1.2 Numpy 数组运算库

NumPy 是使用 Python 进行科学计算的基础包。它包含以下特点：以一个强大的 N 维数组作为研究对象、拥有复杂的（广播）功能、可以作为用于集成 C / C++ 和 Fortran 代码的工具、含有线性代数和傅里叶变换以及随机数功能。

除了明显的科学用途外，NumPy 还可以用作通用数据的高效多维容器。可以定义任意数据类型。这使 NumPy 能够无缝快速地与各种数据库集成。

在需要安装以及使用 Numpy 时，操作者只需要在终端执行：

```
“pip install numpy”
```

等待安装完成之后，在需要使用 Numpy 的 Python 文件中，操作者只需要在文件头部输入一行：

```
“import numpy as np”
```

即可以在代码中利用 “np+.+功能” 来使用 Numpy 中的功能。

2.1.3 Pandas 数据处理库

Pandas 是一个开源的，拥有 BSD 许可的库，为 Python 编程语言提供高性能，易于使用的数据结构和数据分析工具。其旨在使“关系”或“标记”数据的使用既简单又直观，并旨在成为在 Python 中进行实际的、真实的数据分析的基础高级构建块。此外，它还有更广泛的目标，即成为任何语言中最强大，最灵活的开源数据分析/操作工具。它已朝着这一目标迈进。

Pandas 拥有以下的优点：可以轻松的处理浮点中的缺失数据（表示为 NaN）以及非浮点数据、以 DataFrame 为研究对象的大小可变性、强大的兼容性、直观的合并和连接数据集、灵活的数据集整形和旋转、拥有强大的 IO 工具等。

在需要安装以及使用 Pandas 时，操作者只需要在终端执行：

```
“pip install pandas”
```

等待安装完成之后，在需要使用 Pandas 的 Python 文件中，操作者只需要在文件头部输入一行：

```
“import pandas as pd”
```

即可以在代码中利用 “pd+.+功能” 来使用 Pandas 中的功能。

2.1.4 Statsmodels 统计模型库

Statsmodels 是一个 Python 包，为 SciPy 提供了包括描述性统计和统计模型的估计和推断的统计模块的补充。它是建立在 NumPy 和 SciPy 两个库之上的，与 Pandas 集成进行数据处理，为其他的 Python 库提供统计后端。

statsmodels.tsa 包含对时间序列分析有用的模型类和函数。基本模型包括单变量自回归模型（AR），向量自回归模型（VAR）和单变量自回归移动平均模型（ARMA）。非线性模型包括马尔可夫切换动态回归和自回归。它还包括时间序列的描述性统计，例如自相关，部分自相关函数和周期图，以及 ARMA 或相关过程的相应理论属性。它还包括使用自回归和移动平均滞后多项式的方法。此外，还提供相关的统计测试和一些有用的辅助函数。

使用的 statsmodels.tsa.seasonal. seasonal_decompose() 函数的作用是使用移动平均线进行季节性分解，将原始序列分解为趋势、周期、残差三部分。即原始时间序列为 $Y[t]$ ，则利用该函数将 $Y[t]$ 分解为：

$$Y[t] = T[t] + S[t] + R[t] \quad (2-1)$$

其中， $T[t]$ 表示分解出来的趋势（trend）部分， $S[t]$ 表示分解出来的周期（seasonal）部分， $R[t]$ 表示分解出来的残差（residual）部分。

2.1.5 Hmmlern 机器学习库

Hmmlern 是一个开源的，商业上拥有 BSD 许可证的，用于利用 Python 语言来学习 HMM（隐马尔可夫模型）的工具库。其中包含简单算法和模型，尽可能的遵循 scikit-learn API，但适应于序列数据，它建立在 scikit-learn，NumPy，SciPy 和 matplotlib 之上。

hmmlern 中包含有三种隐马尔可夫模型：1）GaussianHMM。该模型中的观测状态是连续状态，且符合高斯分布。2）GMMHMM。该模型中的观测状态是连续状态，且符合混合高斯分布。3）MultinomialHMM。该模型中的观测状态是离散的。

在需要安装以及使用 Hmmlern 时，操作者只需要在终端执行：

“pip install hmmlern”

等待安装完成之后，在需要使用 hmmlern 的 Python 文件中，操作者只需要在文件头部输入一行：

“import hmmlern”

即可以在代码中利用 “hmmlern ++ 功能” 来使用 hmmlern 中的功能。

2.1.6 TensorFlow 深度学习框架

TensorFlow 是一个使用数据流图进行数值计算、用于机器学习的端到端开源平台。在其中图形的节点用来表示数学运算，而图形的边缘用来表示在它们之间流动的多维数据阵列（张量）。这种灵活的体系结构使得操作者可以将计算部署到桌面、服务器或移动设备中的一个或多个 CPU 或 GPU 上而无需重新写代码。TensorFlow 还包括 TensorBoard，一种数据可视化工具包。

TensorFlow 最初是由在 Google 机器智能研究组织的 Google Brain 团队中的研究人员和工程师开发的，开发其目的是更方便的进行机器学习和深度神经网络研究。TensorFlow 的通用性足以适用于各种其他领域。

TensorFlow 为 Python 和 C 语言提供稳定的 API 以及为 C ++, Go, Java, JavaScript 和 Swift 等语言提供非保证的向后兼容的 API。同时 TensorFlow 拥有全面、灵活的工具库和社区资源生态系统，这可让研究人员推动机器学习的最新技术，开发人员可轻松构建和部署机器学习驱动的应用程序。

在安装 TensorFlow 时，本文的实验环境是使用的苹果电脑 MacBook Pro，其显卡型号为 NVIDIA GeForce GT 750M 2 GB 与 Intel Iris Pro 1536 MB，其中 NVIDIA GeForce GT 750M 2 GB 为独立显卡。因此我们可以使用 CPU 版本的 TensorFlow 来进行实验，也可以使用 GPU 版本的 TensorFlow 来进行实验。在这里本文由于使用的数据量并不是特别大，因此最终选取 CPU 版本的 TensorFlow 来进行实验。由于本次实验是在有 Anaconda 环境的电脑中进行安装，因此只需在终端输入“`pip3 install tensorflow`”，若执行不成功则输入“`sudo pip install --upgrade tfBinaryURL`”。

在使用 TensorFlow 时一般在 Python 文件的头部输入“`import tensorflow as tf`”来调用 TensorFlow。在利用 TensorFlow 进行实验时，利用 TensorFlow 中的节点传入数据，调用 TensorFlow 中的结构如神经网络等对数据进行运算。其中 TensorFlow 中的节点包含 `constant`、`placeholder` 以及 `Variable`，`constant` 是 TensorFlow 的常量节点，通过 `constant` 方法创建，其是计算图（Computational Graph）中的起始节点，是传入数据；`placeholder` 是 TensorFlow 的占位符节点，由 `placeholder` 方法创建，其也是一种常量，但是由用户在调用 `run` 方法是传递的，也可以将 `placeholder` 理解为一种形参。即其不像 `constant` 那样直接可以使用，需要用户传递常数值；`Variable` 是 TensorFlow 的变量节点，通过 `Variable` 方法创建，并且需要传递初始值。在使用前需要通过 TensorFlow 的初始化方法进行初始化。

2.2 传统时间序列预测方法

在深度学习神经网络应用于交通流量预测之前,使用传统的基于统计学的交通流量预测方法一直是交通流量预测领域的核心方法。传统的交通流量预测方法多以统计学规律为基础,主要遵循时间序列预测方法的关键和特点,即假设事物的发展趋势会延伸到未来,且预测所依据的数据具有不规则性,还不考虑事物发展之间的因果关系。下面会介绍一些传统的基于时间序列预测方法的交通流量预测方法,如 AR、VAR、MA、ARMA 以及其变体 ARIMA 等。

2.2.1 自回归模型

自回归 (Auto-Regressive, AR) 模型它是一种呈线性的预测模型,它用于描述自然界,经济学等中的某些时变过程。自回归模型指定输出变量线性地依赖于其自己的先前值和随机项 (不完全可预测的术语),因此在机器学习中,该模型采用随机差分方程的形式。利用已知数据,自回归模型从一系列定时步骤中学习,并将先前动作的测量值作为回归模型的输入,以便预测下一个时间步长的值,可以根据 AR 模型来推断出已知数据最末尾数据值的前方或者是后方的数据。AR 模型的主要目的是增加数据量,提高有效可用数据的数量,这与插值法十分类似,而插值法是根据两点 (少数几点) 去推断多点,AR 模型是由已知的多数点进行递推从而得到插入数据,由于 AR 模型初始数据量的规模较大,所以相比之下 AR 模型的效果会好于插值法的效果。

符号 AR (p) 表示订单 p 的自回归模型。AR (p) 模型定义为:

$$X_t = c + \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t \quad (2-2)$$

其中 $\varphi_1, \dots, \varphi_p$ 是模型的参数, c 是一个常数, 而 ε_t 是白噪声。

2.2.2 向量自回归模型

向量自回归 (Vector Auto-Regressive, VAR) 模型是一种常用的计量经济模型,它用模型中所有当期变量对所有变量的若干滞后变量进行回归。向量自回归 (VAR) 是一种随机过程模型,用于捕获多个时间序列之间的线性相互依赖关系。VAR 模型通过允许多个演化变量来推广单变量自回归模型 (AR 模型)。VAR 中的所有变量以相同的方式进入模型: 每个变量都有一个方程式,根据自己的滞后值,其他模型变量的滞后值和误差项来解释其演化。VAR 建模不需要像使用联立方程的结构模型那样需要关于影响变量的力的知识: 所需的唯一先验知识是可

以假设在跨期间相互影响的变量列表。

VAR 模型描述了在相同的采样周期 ($t=1, \dots, T$) 中一组 k 个变量 (称为内生变量) 的演变, 它们只是其过去值的线性函数。变量被收集在一个由 k 组成的 ($k \times 1$) 维矩阵—向量 y_t 中, 其中 $y_{i,t}$ 表示第 i 个变量在时间 t 处的观察值。

2.2.3 移动平均模型

在时间序列分析中, 移动平均 (Moving Average, MA) 模型是用于单变量时间序列建模的常用方法。移动平均模型指定输出变量线性地取决于随机 (不完全可预测) 项的当前和过去值。

MA (q) 是指 q 阶的移动平均模型:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2-3)$$

其中, μ 是序列的平均值, $\theta_1, \dots, \theta_q$ 是模型的参数, $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ 是白噪声误差项, q 被称为 MA 模型的阶数。

2.2.4 自回归滑动平均模型

在时间序列的统计分析中, 自回归滑动平均 (Auto-Regressive and Moving Average, ARMA) 模型根据两个多项式提供 (弱) 静态随机过程的简约描述, 一个用于自回归 (AR), 第二个用于移动平均 (MA)。一般的 ARMA 模型描述于 1951 年 Peter Whittle 的论文: 时间序列分析中的假设检验, 并在 1970 年由 George E. P. Box 和 Gwilym Jenkins 的书中推广。

给定时间序列的数据 X_t , ARMA 模型是理解并且可能预测本序列中未来值的工具。AR 部分涉及对其自身滞后 (即过去) 值的变量进行回归。MA 部分涉及将误差项建模为同时发生并且在过去的不同时间发生的误差项的线性组合。该模型通常被称为 ARMA (p, q) 模型, 其中 p 是 AR 部分的阶数, q 是 MA 部分的阶数。

ARMA (p, q) 是指具有 p 个自回归项和 q 个移动平均项的模型。该模型包含 AR (p) 和 MA (q) 模型:

$$X_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2-4)$$

2.2.5 差分整合移动平均自回归模型

在统计和计量经济学中, 特别是在时间序列分析中, 差分整合移动平均自回

归 (Auto-Regressive Integrated Moving Average, ARIMA) 模型是自回归移动平均 (ARMA) 模型的推广。ARIMA 模型适用于数据显示非平稳性证据的某些情况, 其中初始差分步骤 (对应于模型的“整合”部分) 可以应用一次或多次以消除非平稳性。

ARIMA 中的 AR 部分表明, 进化的感兴趣变量在其自身的滞后 (即先前) 值上回归。MA 部分表示回归误差实际上是误差项的线性组合, 其值在同一时间和过去的不同时间发生。I 表示数据值已被其值与先前值之间的差异替换 (并且该差异过程可能已执行多次)。这些功能中的每一个的目的是使模型尽可能地适合数据。

非季节性 ARIMA 模型通常表示为 ARIMA (p, d, q), 其中参数 p, d 和 q 是非负整数, p 是自回归模型的阶数 (时间滞后数), d 是差分 (数据减去过去值的次数), q 是移动平均模型的顺序。季节性 ARIMA 模型通常表示为 ARIMA(p, d, q) (P, D, Q) m, 其中 m 表示每个季节的周期数, 而大写 P, D, Q 表示自回归, 差分, 和 ARIMA 模型的季节性部分的移动平均项。

给定时间序列数据 X_t , 其中 t 是整数索引而 X_t 是时间序列中的数据值, 一个 ARMA (p', q) 模型可以表示为:

$$X_t - \alpha_1 X_{t-1} - \cdots - \alpha_{p'} X_{t-p'} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (2-5)$$

这等价于:

$$(1 - \sum_{i=1}^{p'} \alpha_i L^i) X_t = (1 + \sum_{i=1}^q \theta_i L^i) \varepsilon_t \quad (2-6)$$

其中, L 是一个滞后因子, α_i 是自回归模型部分的参数, θ_i 是移动平均模型部分的参数, ε_t 是误差项。通常假设误差项 ε_t 是一个独立同分布的从均值为零的正态分布中采样而得到的变量。

现在假设多项式 $(1 - \sum_{i=1}^{p'} \alpha_i L^i)$ 有一个具有多重性 d 的单位根 (一个因子 $(1 - L)$), 那么上述式子可以改写为:

$$(1 - \sum_{i=1}^{p'} \alpha_i L^i) = (1 - \sum_{i=1}^{p'-q} \phi_i L^i) (1 - L)^d \quad (2-7)$$

ARIMA (p, d, q) 过程用 $p = p' - d$ 表示该多项式因式分解性质, 并可以表示为下式:

$$(1 - \sum_{i=1}^p \phi_i L^i) (1 - L)^d X_t = (1 + \sum_{i=1}^q \theta_i L^i) \varepsilon_t \quad (2-8)$$

因此可以认为 ARIMA (p, d, q) 过程是具有 d 个单位根是自回归多项式 ARMA (p + d, q) 过程的特定情况。因此, 当 $d > 0$ 是, ARIMA (p, d, q) 过程不会是一个广义平稳过程。

以上的推导可以概括如下:

$$(1 - \sum_{i=1}^p \phi_i L^i) (1 - L)^d X_t = \delta + (1 + \sum_{i=1}^q \theta_i L^i) \varepsilon_t \quad (2-9)$$

上式表示带有偏移 $\delta / (1 - \sum \phi_i)$ 的 ARIMA (p, d, q) 过程。

2.3 机器学习算法

机器学习是一门科学，它基于对算法和统计模型的科学研究，依赖模型和推导，不使用明确的指令和要求，通过计算机系统使用它来有效的执行一些特定的任务。机器学习是人工智能领域中的一个分支，并在业界内被认为是最可以体现智能化的一个部分。机器学习算法通过构建样本数据的数学模型，在没有明确的编程任务的情况下进行预测或者决策。

根据样本数据（训练数据）是否拥有标记信息，机器学习任务大致可分为四大类别：有监督学习、无监督学习、半监督学习和强化学习。其中主要包括分类算法、回归算法、聚类算法等。

这些机器学习算法的应用十分的广泛，它不仅仅在计算机科学中的诸多领域中占有独特的地位，如图像处理、数据挖掘、自然语言处理等，还为许多交叉学科的发展提供了强大的支持，如借用机器学习对相应的数据进行系统的处理，为该学科的深入研究奠定了坚实的基础。

2.3.1 隐马尔可夫模型

隐马尔可夫模型（Hidden Markov Model, HMM）是一个含有未观察（即隐藏）状态的马尔可夫过程的统计马尔可夫模型。它可以表示最简单的动态贝叶斯网络，与早期关于最优非线性滤波问题的研究密切相关。隐马尔可夫模型是一种经典的有向图模型，应常被用于时序问题的分析。

隐马尔可夫模型是马尔可夫模型的一种延伸。在更为简单的马尔可夫模型（如马尔可夫链）中，状态可以被观察到，即观察者可见状态，因此马尔可夫链的唯一参数是状态转移概率。而在隐马尔可夫模型中，状态是一个隐变量，并不是直接可见的，可是依赖于该状态下的以数据或其他隐藏格式（token）表现的输出是可见的，每一个状态都有一个相应的可能输出的概率分布。因此由 HMM 生成的结果序列给出了关于状态序列的一些信息，这也被称为模式理论。隐马尔可夫模型中的“隐”指的是隐藏的状态序列而不是模型的参数，即使模型的参数是已知的，由于不可知的隐藏状态序列，这个模型依旧被称为隐马尔可夫模型。

图 2-1 给出了通用的隐马尔可夫模型的结构，其中 $x(t)$ 表示隐藏状态序列， T 是观察序列的长度， $O(i)$ 是观测序列。

A 是隐含状态转移概率，表示隐马尔可夫模型中各个状态之间的转移概率， A 值构成的矩阵为隐含状态转移矩阵， $A=[a(ij)]^{N \times N}$ ，其中： $a(ij)=P(x(t+1)=s(j)|x(t)=s(i))$ ， $1 \leq i, j \leq N$ 表示在任一时刻 t 时该状态为 $s(i)$ ，则下

一时刻 $(t+1)$ 时的状态为 $s(j)$ 的概率。

B 是观察状态转移概率, 表示隐马尔可夫模型根据当前的状态得到各个观察状态的概率, B 值构成的矩阵为观察状态转移矩阵, $B=[b(ij)]^{N \times M}$, 其中: $b(ij)=P(y(t)=O(j)|x(t)=s(i))$, $1 \leq i \leq N$, $1 \leq j \leq M$ 。表示在任一时刻 t 时该状态为 $s(i)$ 其观测值 $O(j)$ 被获取的概率。

π 是初始状态概率, 表示在开始时模型中各个隐状态出现的概率, $\pi=(\pi_1, \pi_2, \dots, \pi_N)$, 其中: $\pi_i=P(x(0)=s(i))$ 。表示该隐马尔可夫模型的初始状态为 s_i 的概率。

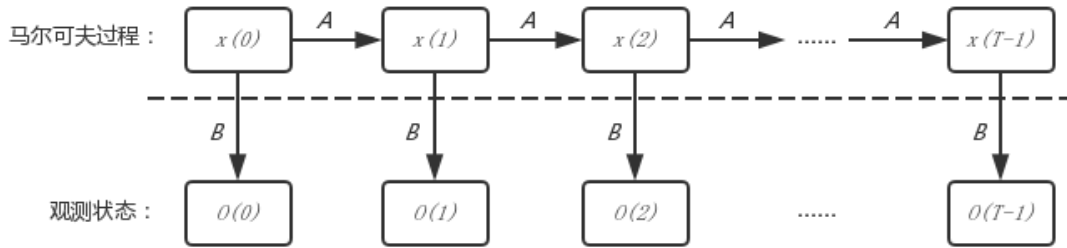


图 2-1 隐马尔可夫模型
Figure 2-1 Hidden Markov Model

根据上述三个参数以及观测空间和状态空间, 就可以确定一个隐马尔可夫模型, 该隐马尔可夫模型用其参数 $\lambda=[A, B, \pi]$ 来表示。

根据图 2-1 可以看出由给定隐藏变量 x 的值得到的隐藏变量 $x(t)$ 在时间 t 的条件概率分布, 仅取决于隐藏变量 $x(t-1)$ 的值; 而与时间 $t-2$ 和其之前的值没有关系。这被称为马尔可夫过程, 即系统的下一时刻的结果状态只由当前的状态决定, 而与以往的其他状态无关。类似地, 观测变量 $y(t)$ 的值仅取决于隐藏变量 $x(t)$ 的值 (两者在时间 t)。

2.4 深度学习算法

深度学习是机器学习的一个分支, 是基于学习数据的表征不是特定于任务的算法的更广泛的机器学习方法, 也被叫做深度结构化学习或者分层学习。现今大多数深度学习模型都基于人工神经网络 (ANN), 这是一种受动物大脑的生物神经网络启发的计算系统, 通过模仿大脑中神经元之间的信号传递来实现信息的处理。

图 2-2 展示了神经元的状态与单层神经网络的结构。神经元的数学表达式为:

$$y = f(\sum_{i=1}^n \omega_i x_i - \theta) \quad (2-10)$$

其中 x_i 是第 i 个神经元的输入, ω_i 是第 i 个神经元的连接权重, θ 是当前神经元的

阈值， f 是当前神经元的激活函数。当前神经元收到了来源于它前面 n 个神经元输出的传导过来的输入信号值，这些信号值根据发出信号的神经元与当前神经元的不同链接权重对该信号进行放大缩小，将进行缩放后的结果进行加权求和并与该神经元的阈值相比较，将比较后的结果经过“激活函数”处理就得到了当前神经元的输出。

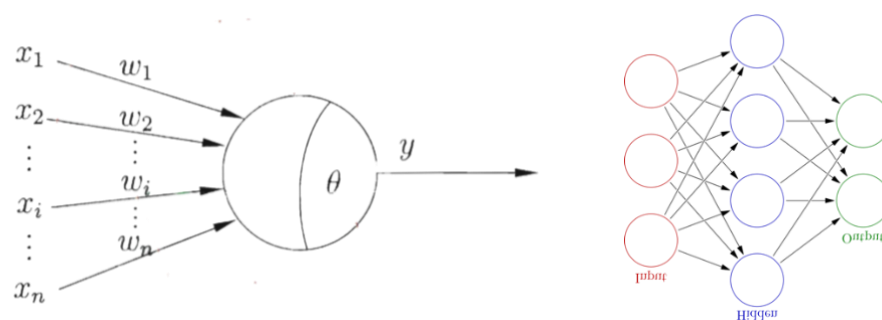


图 2-2 神经元状态及神经网络结构
Figure 2-2 Neuron Status and Neural Network Structure

深度神经网络（DNN）是人工神经网络的一种，在输入层和输出层之间有多层结构，无论是输入与输出是线性关系还是非线性关系，深度神经网络会找到一种正确的数学运算方法使得输出能由输入得到。

深度神经网络（DNN）可以模拟复杂的非线性关系，它通过堆叠多层结构，组合低层的特征，得到更加抽象的高层特征或者类别属性表示，以发现数据的更多特征。现阶段主流的深度学习算法主要为循环神经网络（RNN）、卷积神经网络（CNN），其中循环神经网络（RNN）主要应用于序列数据的处理，如机器翻译、语音识别、序列预测等；卷积神经网络（CNN）主要应用在计算机视觉领域，如图像处理、图像识别、图像分类等。接下来将介绍这两种主要的深度学习算法并介绍一种卷积神经网络（CNN）的变体—图卷积神经网络（GCN）。

2.4.1 循环神经网络

传统的神经网络已经可以处理很多的工作，在其工作过程中，输入数据经由输出层到隐藏层最终到达输出层，在这个过程中所有的层内神经元之间是不连通的，而所有的层与层之间都是全连接（FC）的，即同层神经网络内各个神经元数据无法进行交互，而相邻层之间任意两个神经元是可以连通的。这种特殊的结构导致在处理某些问题上很棘手，因为在同一神经网络层中，各个神经元相互独立，但若有些数据在相邻神经元上有一定的联系，这将无法体现。如在根据上文预测下文的文字时，需要用到上文的内容以此来推测下文的内容，因为文章的前后文并不是完全独立的而是有一定联系的。这时就需要用到循环神经网络（RNN）

了。

RNN 之所以称为循环神经网络，实际上当前神经元的输出不仅仅取决于当前神经元的输入，还取决于以往神经元的输出。其内部的具体原理是网络会对输出信息进行存储并作用于其后面的神经元，体现在网络结构上的变化是在隐藏层中各个神经元不再独立而是之间存在连接，隐藏层神经元的输入也不仅仅是输入层的输出，还包涵上一时刻隐藏层神经元的输出。

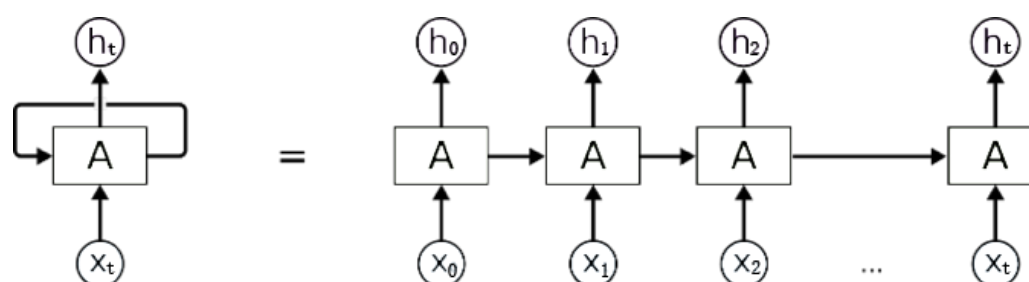


图 2-3 循环神经网络
Figure 2-3 Recurrent Neural Network

图 2-3 为 RNN 的示例图，其中 A 为循环神经网络中的神经元， x_i 为输入层的输出值（即隐藏层的输入值）， h_i 为循环神经网络层的输出值（即隐藏层的输出值）。由图中的循环结构可以看出当前的神经元的输出结果可以通过该循环结构传递到下一神经元，与下一神经元得到的输入层输出值一同作为下一神经元的输入值。图 2-3 右侧为图 2-3 左侧的展开形式，从图 2-3 右侧中可以更好的理解神经元中的消息传递，这种链式特征更好的揭示了循环神经网络本质上与列表和序列是有很大的相关性的。

虽然上述循环神经网络可以解决输入数据之间有隐含联系的问题，但是针对这其中的不同问题，RNN 的效果有可能会有很大的差别。例如，在进行上下文推测的时候，有的预测下文与很相近的上文有较大联系，即预测依赖的相关信息与预测信息的位置之间间隔较小，位置较为相近，这样 RNN 的预测效果较好，而当预测下文与相隔较远的上文有较大联系时，即预测依赖的相关信息与预测信息的位置之间间隔较大，位置相对较远，这样 RNN 的预测效果就会变差。理论上 RNN 可以处理无限长度的长期依赖问题，但在实验中发现，RNN 并不能准确的学习到依赖信息较为久远的序列，预测结果不尽如人意。因此有人提出了长短期记忆（Long Short-Term Memory, LSTM）网络。

长短期记忆网络（LSTM）是在 1997 年被提出的，它主要为解决长期依赖问题而设计，它一经提出就占领了多个领域的预测准确率排行榜第一名的位置。LSTM 未改变 RNN 的外部整体结构，而是对 RNN 中神经元的内部结构进行调整。RNN 的内部结构如图 2-4 所示，在传统的标准 RNN 中，重复的神经元模块只含有一个非常简单的结构，将隐藏层的输入值与上一神经元的输出值进行整合，

把整合后的结果送入一个激活函数（例如 $\tanh(\cdot)$ ），得到的输出即为当前神经元的输出。其数学表达式为：

$$h_t = \sigma_h(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \quad (2-11)$$

其中， W_{xh} 是输入层的输出到隐藏层的权重矩阵， W_{hh} 是隐藏层两个神经元之间的权重矩阵， b_h 是隐藏层的偏置向量， σ_h 是隐藏层的激活函数。

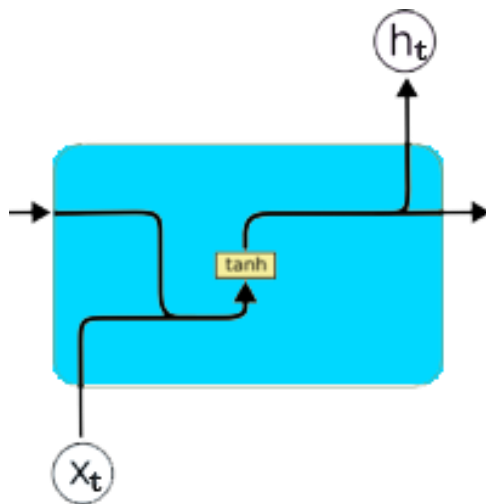


图 2-4 RNN 内部结构

Figure 2-4 Internal Structure of RNN

LSTM 的外部组成结构与 RNN 结构相似，但是在重复的神经元模块中拥有不同的内容。LSTM 中最重要的就是 LSTM 中的神经元的状态，即细胞状态（cell state）。在 LSTM 中存在三个门结构，即遗忘门、输入门、输出门。利用细胞状态以及这三个门结构的共同作用，就可以实现 LSTM 的功能。LSTM 的结构如图 2-5 所示。其中， h_{t-1} 是上一个神经元结构的输出， x_t 是当前神经元结构的输入。

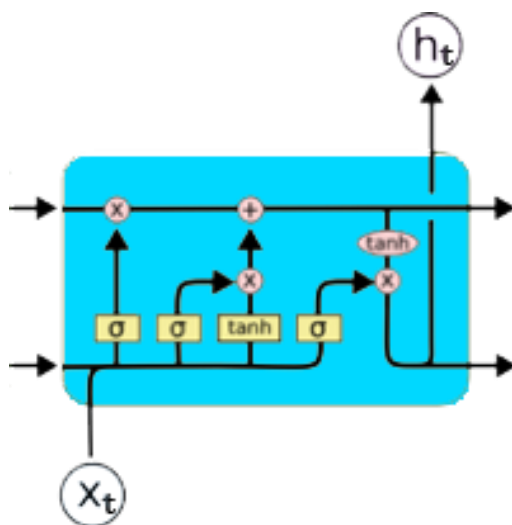


图 2-5 LSTM 内部结构

Figure 2-5 Internal Structure of LSTM

遗忘门决定了在以前的细胞状态中要丢弃什么信息，其数学表达式为：

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2-12)$$

输入门决定了要将多少新的信息加入到细胞状态之中，其数学表达式为：

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2-13)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2-14)$$

输出门决定了当前的神经元结构要输出什么值，其数学表达式为：

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2-15)$$

$$h_t = o_t * \tanh(C_t) \quad (2-16)$$

2.4.2 卷积神经网络

卷积神经网络（CNN）是一种专门处理网格类数据的深度神经网络，利用多层感知机的变体设计而成，常被用于计算机视觉领域，主要体现在图像与视频识别、图像分类、图像分析等方面。与传统的图像处理算法（如 SIFT）想比，CNN 在数据预处理工作上所耗费的时间与资源更少，原因是 CNN 利用深度神经网络来设计传统图像处理算法中的滤波器。与传统神经网络相比，CNN 利用输入是图片的特点，将神经元设计成三维立体的状态，具体结构如图 2-6 所示。

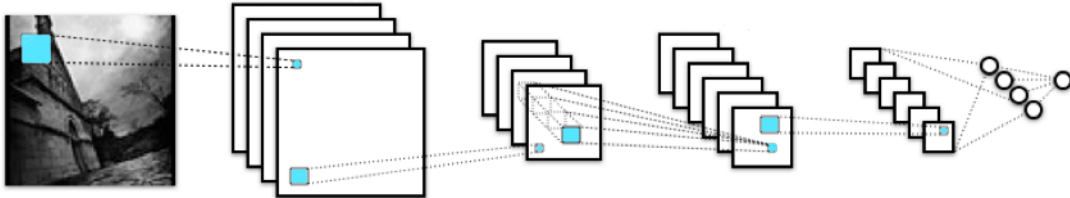


图 2-6 卷积神经网络

Figure 2-6 Convolutional neural network

CNN 通常包括以下几种层结构：

卷积层：利用卷积核进行卷积运算，提取输入的不同特征，低层提取较为低级的特征，高层将提取出的低级特征进行迭代从而从其中提取出更加复杂的特征。

线性整流层：使用线性整流最为激活函数增强网络的非线性特征。

池化层：实质上是在进行降采样，将输入图像划分成一组非重叠矩形，并且输出每个子区域中的最大值。

全连接层：将所有的局部特征整合成为全局特征。

2.4.3 图卷积网络

本文在上一小节介绍了 CNN，其中卷积的本质是利用一个共享参数的卷积核，通过计算中心像素点以及相邻像素点的加权和来构成 feature map 实现空间特征的提取，其中的加权系数就是卷积核的权重系数。而 CNN 适用的场景之前也提到了，其处理的图像或者视频数据中像素点 (pixel) 是排列成很整齐的矩阵，即具有欧几里得结构 (Euclidean Structure)。而在现实的科学研究中还有很多非欧几里得结构 (Non Euclidean Structure) 的数据，如社交网络、信息网络等，这类图结构就无法使用 CNN 来进行处理，因此提出了图卷积网络 (GCN)，这里的图是指数学 (图论) 中的用顶点和边建立相应关系的拓扑图。

与 CNN 的主要功能--提取图片的特征信息一样，GCN 是用来提取拓扑图的空间信息的。CNN 利用卷积核来实现卷积操作，而在 GCN 中其借用图谱的理论来实现拓扑图上的卷积操作，图谱理论是借助于图的拉普拉斯矩阵的特征值和特征向量来研究图的性质。考虑图上的频谱卷积定义为信号 $x \in \mathbb{R}^N$ (每个节点的标量) 与傅立叶域中由 $\theta \in \mathbb{R}^N$ 参数化的滤波器 $g_\theta = \text{diag}(\theta)$ 的乘法，谱图卷积的定义为：

$$g_\theta \star x = U g_\theta U^T x \quad (2-17)$$

其中， U 是归一化后的图拉普拉斯 $L = I_N - D^{-1/2} A D^{-1/2} = U \Lambda U^T$ 特征向量矩阵， L 是图 G 的拉普拉斯算子， A 是图 G 的邻接矩阵，图的特征值对角矩阵 Λ 和 $U^T x$ 共同构成了 x 的图傅立叶变换。可以将 g_θ 理解为一个 L 的特征值的函数，即 $g_\theta(\Lambda)$ 。由于公式(2-17)的计算量是非常庞大的，因此 Hammond^[41]等人提出了 $g_\theta(\Lambda)$ 可以通过切比雪夫多项式 $T_k(x)$ 的 K 阶截断展开很好的近似出来：

$$g_{\theta'}(\Lambda) \approx \sum_{k=0}^K \theta'_k T_k(\tilde{\Lambda}) \quad (2-18)$$

其中， $\tilde{\Lambda} = (2/\lambda_{\max})\Lambda - I_N$ ， λ_{\max} 是 L 的最大特征值， $\theta' \in \mathbb{R}^K$ 是切比雪夫系数的向量。切比雪夫多项式递归的定义为 $T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)$ ，其中 $T_0(x) = 1$ 、 $T_1(x) = x$ 。

根据公式(2-17)与公式(2-18)，可以得到：

$$g_{\theta'} \star x \approx \sum_{k=0}^K \theta'_k T_k(\tilde{L})x \quad (2-19)$$

其中， $\tilde{L} = (2/\lambda_{\max})L - I_N$ 。很容易知道 $(U \Lambda U^T)^k = U \Lambda^k U^T$ 。但是要注意到公式(2-19)是 K 阶局部的，因为公式(2-19)是拉普拉斯算子的 K 阶多项式，即它仅仅取决于距离中心节点最大 K 步的节点 (中心节点的 K 阶邻域)。Defferrard^[42]等人使用这个 K 局部化卷积来定义图上的卷积神经网络。

因此，可以通过堆叠多个式**的形式的卷积层结构来构建图卷积神经网络模型。现在假设将卷积运算限制为 $K=1$ ，即变为是图拉普拉斯谱的线性函数。进一步近似 $\lambda_{\max} \approx 2$ ，则公式(2-29)简化为

$$g_{\theta'} \star x \approx \theta'_0 x + \theta'_1 (L - I_N) x = \theta'_0 x - \theta'_1 D^{-\frac{1}{2}} A D^{-\frac{1}{2}} x \quad (2-20)$$

其中，具有两个参数 θ'_0 和 θ'_1 ，过滤器参数可以在整个图形上共享。通过连续应用这种形式的滤波器可以有效地卷积中心节点的 k 阶邻域，其中 k 是神经网络模型中的连续滤波操作或卷积层的数量。

在实践中，进一步限制参数的数量以解决过度拟合并最小化每层的操作（例如矩阵乘法）的数量可能是有益的。这给我们留下了以下表达式：

$$g_{\theta'} \star x \approx \theta (I_N + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}) x \quad (2-21)$$

其中 $\theta = \theta'_0 = -\theta'_1$ 。需要注意的是 $I_N + D^{-1/2} A D^{-1/2}$ 的结果是在范围 $[0,2]$ 内，因此当在深度神经网络模型重复中使用 θ 时，导致数值不稳定性以及梯度爆炸或消失。为了解决这个问题，数学家引入了下面这个技巧： $I_N + D^{-1/2} A D^{-1/2} \rightarrow \tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2}$ ，其中 $\tilde{A} = A + I_N$ ， $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ 。

我们可以将该定义推广到具有 C 个输入通道的信号 $X \in \mathbb{R}^{N \times C}$ （即每个节点的 C 维特征向量）和 F 滤波器或特征图上，如下所示：

$$Z = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} X \Theta \quad (2-22)$$

其中 $\Theta \in \mathbb{R}^{C \times F}$ 是滤波器参数矩阵， $Z \in \mathbb{R}^{N \times F}$ 是卷积信号矩阵。

因此最终考虑使用具有以下分层传播规则的多层图形卷积网络：

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}) \quad (2-23)$$

其中， $\tilde{A} = A + I_N$ 是加入自链接的无向图 G 的邻接矩阵， I_N 是单位矩阵， $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ ， $W^{(l)}$ 是训练得出的层权重矩阵， $\sigma(\cdot)$ 是激活函数，例如 $\text{ReLU}(\cdot) = \max(0, \cdot)$ 。 $H^{(l)} \in \mathbb{R}^{N \times D}$ 是第 l 层的激活矩阵，其初始化 $H^{(0)} = X$ 。

2.5 线图

线图与大家通常所说的折线图是不同的概念，在这里线图是一个图论概念。一个无向图 G 的线图 $L(G)$ 表现了图 G 中各条边之间的邻接关系，给定图 G ，其线图 $L(G)$ 定义如下：

$$V(L(G)) = E(G) \quad (2-24)$$

$$E(G) = \{(e, e') | e \in G, e' \in G, e \cap e' = V \in G\} \quad (2-25)$$

线图 $L(G)$ 的每个顶点表示图 G 中的一条边，线图 $L(G)$ 中两个顶点之间有连边当且仅当这两个顶点在图 G 中所对应的连边共享一个顶点。

通俗解释无向图 G 与其线图 $L(G)$ 的关系就是将图 G 中的顶点转换为 $L(G)$ 中的连边，将图 G 中的连边转换为 $L(G)$ 中的顶点，在 $L(G)$ 中顶点表示图 G 中的连边， $L(G)$ 中的连边表示该连边连接的两个顶点在图 G 中对应的连边是相连的，实现过程如图 2-7 所示。

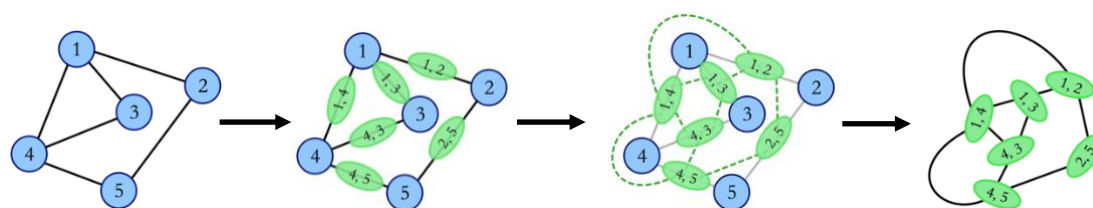


图 2-7 线图

Figure 2-7 Line Graph

2.6 本章小结

本章主要对实验平台及相关理论知识进行了介绍，主要描述了本文实验的技术平台以及相应的模块搭建。主要介绍了 Python 以及与数据分析、挖掘相关的扩展包：Numpy、Pandas、Statsmodels、hmmlearn，接下来介绍了深度学习平台 TensorFlow 的原理、安装及其应用。下面介绍了常用的时间序列预测方法，其中包含传统的基于统计学的交通流量预测方法如 ARIMA 等。接下来介绍了机器学习的概念及发展，并阐述了机器学习中常用于时间序列分析的隐马尔可夫模型（HMM 模型）。之后在机器学习的基础上探究了深度学习的定义和深度神经网络的应用，并介绍了两种常用的深度神经网络模型：卷积神经网络（CNN）和循环神经网络，并在此基础上介绍了图卷积神经网络，对这三个模型所应用的场景、所解决的问题及相应的原理进行了细致的阐述。最后介绍了本文所涉及的一种图论中数学定义——线图，描述了其功能及含义，并对其概念进行了通俗的解释。本章对全文所涉及的概念、原理及框架进行了系统的梳理，对读者们对本文的理解提供了基础知识的支持，让读者们能够更好对本文的内容进行理解。

3 系统设计与数据分析处理

本章主要对本文所提出的流量预测系统的整体框架进行具体论述,同时对所处理数据进行描述与分析。首先根据本文的需求对交通流量预测时应注意的问题进行梳理,之后根据这些问题本文设计了相应的处理流程,包括数据解析与模型算法两方面。接下来对数据集进行观察与分析,通过统计学观察数据的相应特征关系,并对数据进行预处理。最后对本章内容进行总结。接下来,本章将主要从系统设计、数据分析和数据预处理三方面进行具体说明。

3.1 系统设计

在构造交通流量预测系统时,本文要根据交通流量的特征来对系统进行设计。在交通流量预测问题中,要关注的有以下几点:

(1) 在研究道路上的交通流量情况时,相对于选取道路交叉路口处的车辆数据而言,更应选取道路上的车辆数量值转化为交通流量作为研究对象,这样能够更明显的体现道路上的交通情况。

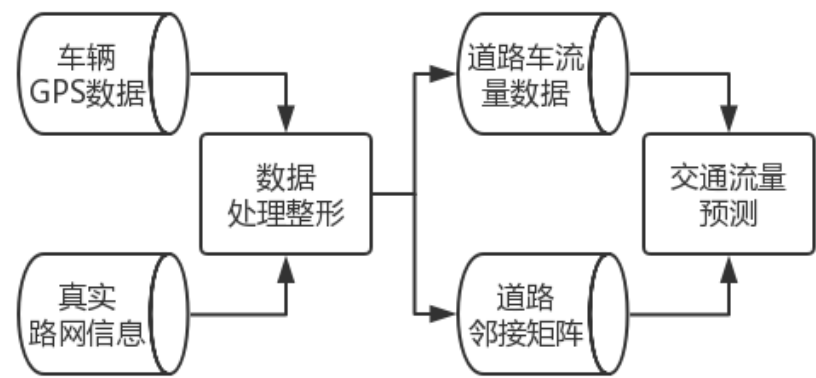
(2) 在分析道路上的交通流量特征时,不能像以往的方法一样仅着眼于时间维度,而是要同时兼顾时间维度与空间维度两方面,共同作用来分析流量数据,而且在空间维度上还要遵循实际的路网结构。这体现在实际中就是不仅仅要着眼于所选道路的以往流量值,还要分析与该道路相关道路上的流量值。

(3) 在时间、空间维度之外,本文还认为对于同一路段而言,每天的交通流量变化是会成周期性的,因此还需要引入时间周期特征。除此之外,还有一些其他特征,如天气等因素也有可能对道路上的流量值产生一定的影响,因此也需要同时引入这些其他特征来分析对交通流量预测是否有正向的影响作用。

根据对上述关注点及问题的分析,设计了本文的交通流量预测系统,主要包含两大模块:数据处理模块、流量预测模块。如图 3-1 所示。

在数据处理模块中,包含有数据清洗、数据整合、数据整形等步骤,其中数据清洗是将车辆 GPS 数据中的问题数据删去,留下可用数据;数据整合是对留下的可用数据进行处理,将车辆 GPS 数据整合成道路交通流量数据;数据整形是将道路交通流量数据整形成后面的流量预测模块所需的数据形式。

在流量预测模块中,我们对前面处理后得到的道路交通流量数据进行处理,提取其中的空间特征以及时间特征得到流量预测结果,同时提取长周期特征和其他特征对交通流量进行预测,整合预测结果共同对下一时刻的交通流量进行预测。



3.2 数据分析

本文选用的数据是由某公司提供的北京市城区内的真实出租车 GPS 数据集，由车辆内的控制系统实时上传并汇聚而成。该数据中包含多维度特征，具体情况如表 3-1 所示。

表 3-1 车辆信息数据集字段描述

Table 3-1 Field Description of Vehicle Dataset

中文名称	英文名称	数据类型
SUID 唯一标识 (报文类型)	SUID	VARCHAR
OEM 识别码	OEMCODE	VARCHAR
终端手机 SIM 卡号	COMMADDR	INTGETER
接收报文目标 ID	DST	VARCHAR
UTC 时间	UTC	VARCHAR
报文的长度 (目前尚未使用, 通常置 0)	APC	INTGETER
纬度	LAT	VARCHAR
经度	LON	VARCHAR
行驶角度	HEAD	VARCHAR
行驶速度	SPEED	VARCHAR
里程 (米)	DISTANCE	VARCHAR
定位描述 (定位有效或定位无效)	TFLAG	VARCHAR
空重车描述	VFLAG	VARCHAR
状态 (目前尚未使用)	OST	VARCHAR
状态描述 (车辆当前状态描述)	OSTDESC	VARCHAR

其中，SUID 用来标识该条数据在数据集中的编号，是数据在该车辆数据集中唯一标示；COMMADDR 是上传数据的手机 SIM 卡号，在这里用来标识上传数据的车辆；UTC 用来标识数据采集的时间；LON、LAT 用来标识在数据采集时车辆所在的经纬度。

经过对数据集进行初步统计，对数据集中的车辆数量、时间采集信息、定位是否有效等进行统计整理，结果如表 3-2 所示。

表 3-2 车辆信息统计结果
Table 3-2 Vehicle Information Statistics

统计信息	有关字段	统计结果	单位
车辆总数量	COMMADDR	33043	辆
数据采集速率	UTC	1	条/秒
定位有效	TFLAG	2800000	条/天

接下来选取所有定位有效的数据，根据每条数据中的经纬度将车辆数据点与地图结合，落到路网上，结果如图 3-2 所示。

根据对上述信息的分析，本文将在某一时间段内道路上的车辆经过的数量作为该道路在这一时间段内交通流量。

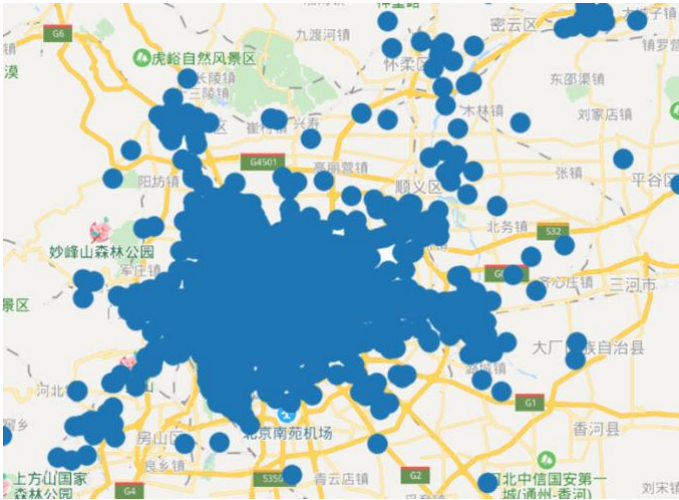


图 3-2 车辆数据分布
Figure 3-2 Vehicle Data Distribution

3.3 数据预处理

根据 3.2 小节中对数据集的分析，对所拥有的数据集进行预处理。预处理的主要工作内容是将车辆 GPS 数据集处理为道路上的交通流量数据。

第一步，将完整车辆 GPS 数据集进行清理，去除数据集中的脏数据以及失

效数据。

第二步，对可用车辆 GPS 数据在时间上进行切分。为使得切分后数据可用，根据其采集速率，选取时间间隔为 10 分钟，将每天的 24 小时切分成 144 个每个长度为 10 分钟的时间片，即每天的交通流量时间序列长度为 144。

第三步，将每个时间片内的车辆 GPS 数据与真实路网数据结合，将车辆 GPS 数据落到路网上，对每一条道路上的车辆信息进行去重，避免重复统计，再对去重后的结果进行数量统计，得到的最后数值将其作为该时间片内这条道路上的交通流量数据。

第四步，将所有处理得到的交通流量数据进行整合，得到每条道路在不同时间片内的交通流量数据构成的流量矩阵。

相应工作流程如图 3-3 所示。

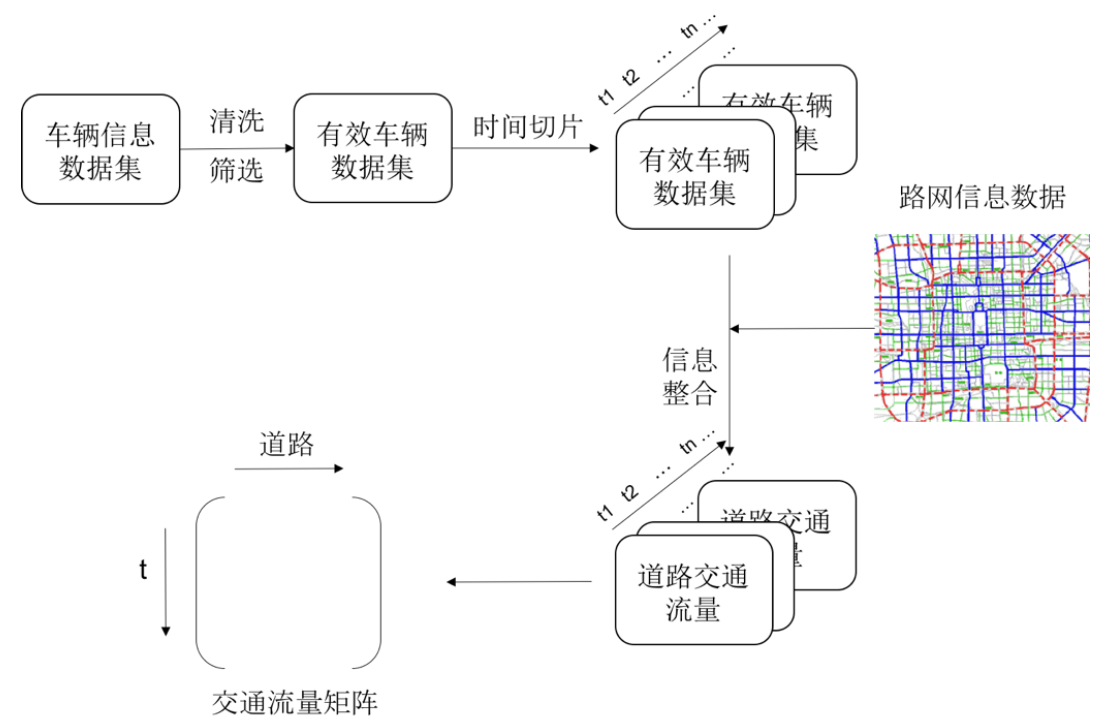


图 3-3 数据预处理工作流程

Figure 3-3 Workflow of Data Preprocessing

接下来对得到的结果进行可视化，用以验证在 3.1 小节中提出的交通流量预测问题所应该关注的要点。

首先观察流量数据的时空关系。选取一条中心道路和与其相连的其他四条道路，在图 3-4 中展示一天内在相应道路上统计得到的流量值的变化。根据图 3-4 可知，某一条道路上的流量值的变化与其上下游道路上流量值的变化有关。

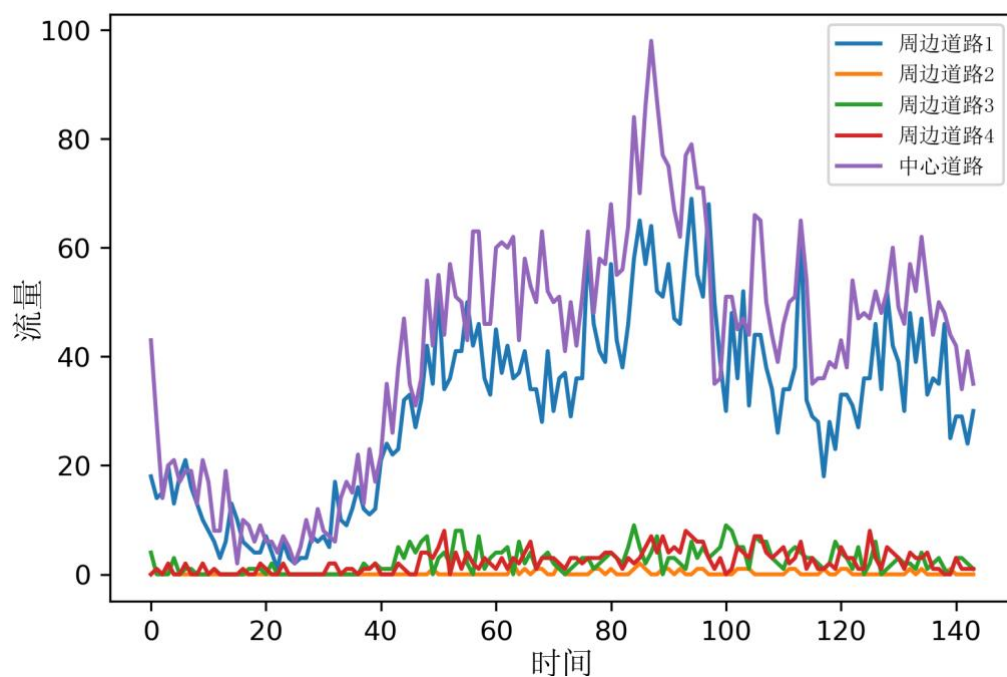


图 3-4 五条相邻道路一天内交通流量变化

Figure 3-4 Traffic Flow Changes in Five Adjacent Roads in One Day

接下来观察道路上流量值变化的周期性。选取一条道路，在图 3-5 中展示在三天内该道路上的交通流量变化。根据图 3-5 可知，某一条道路上的流量值在以天为单位的时间长度上有一定的周期性。

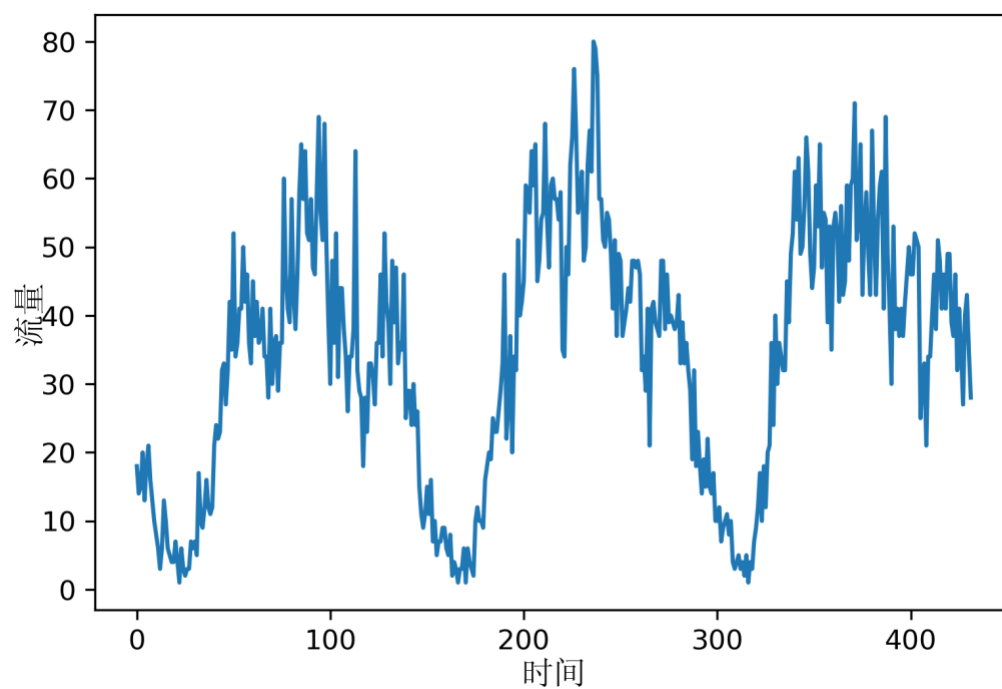


图 3-5 某一条道路上三天交通流量变化

Figure 3-5 Three Days Traffic Flow Changes on a Road

3.4 本章小结

本章主要介绍了交通流量问题所要注意的关键点与问题,并根据这些关注点设计了本文的交通流量预测整体系统,对系统中各个模块的功能进行了简单的描述。同时针对现在所拥有的数据集做了一个分析,详细了解了数据的格式、数据关键字含义以及数据集的相关统计信息,最后对车辆 GPS 数据集进行了基础的预处理操作,明确了数据预处理的工作流程,将数据集中的脏数据去除,并对数据进行切分整形,将数据整理成道路上的交通流量数据,同时对处理后得到的道路交通流量数据进行观察,以此验证在系统设计时本文所提出的需要关注的关键点是必要的。

4 基于深度学习的交通流量预测

本章主要描述了对本文所设计的基于深度学习的交通流量预测系统的实现过程。在介绍本文的基于深度学习的流量预测系统时候，从特征提取方面入手，主要介绍了交通流量的空间特征提取及预测、交通流量的时间特征提取及预测与其他特征提取及预测这三部分内容。最后对所得到的三个预测结果进行整合，得到本文设计的交通流量预测系统的输出。

在特征提取的顺序方面，主要需要判定的是道路上的交通流量数据的空间特征提取与时间特征提取的先后顺序。空间信息的提取是在同一时刻上对不同位置的信息进行提取，而时间信息的提取是在不同时刻上对同一位置的信息进行提取。若先提取时间特征再提取空间特征，则先需要对每一个所选位置的时间特征进行提取，再对所有位置的空间特征进行提取，这样的操作无法实现。因此本文选择先提取道路上的交通流量数据的空间特征，再提取道路上的交通流量数据的时间特征。

4.1 交通流量空间特征提取及预测

空间特征提取，其实就是对数据的空间信息加以描述，并找出空间信息与数据之间存在的联系，提取交通流量数据的空间特征是交通流量预测过程中的关键问题。在大多数情况下，要进行空间特征提取的数据一般都是以一张二维平面图的形式存在的，如地图等，因此只要提取出地图上的空间信息得到的就是隐藏的空间特征。在深度学习与神经网络发展十分迅猛的今天，提取图片的空间特征一般都是使用卷积神经网络（CNN），因为其相对于传统的方法（如 SIFT）来说更为方便、简洁，且效果也足够优秀，所以在一些交通流量数据的使用过程中都使用 CNN 来提取道路交通流量的空间特征。

可是 CNN 提取的一般是完整图片的信息，而车辆 GPS 数据在落到地图上之后虽然可以分块进行统计形成流量矩阵，但是在数据整形过程中就会发现有数值区域相对于空闲区域来说占总区域的比重很小，因为当车辆 GPS 数据与地图匹配时，车辆 GPS 数据将落在地图中的道路上，而道路在地图中的所占比例是很少的，所以造成了没有道路的地方没有数据，统计得到的流量矩阵将有很多为 0 值的地方。而且 CNN 在提取流量矩阵的空间特征时，是利用卷积核扫描矩阵，而这种提取空间特征方式并不符合路网的实际结构。因此，基于上述两个原因本文认为利用 CNN 提取道路交通流量的空间特征是不准确的。

在这里本文则根据路网的实际空间信息提出了本文所用的道路交通流量空间特征的提取方法——利用图卷积网络（GCN）。本文将路网看作拓扑图结构，则道路即为所得到的拓扑图的连边，而道路交点则为所得到的拓扑图的节点。可是这样道路上的交通流量信息都在拓扑图的连边上，与 GCN 处理数据时所要求的输入形式不符，所以在这里本文同时提出了数据转换的方法，即利用图论中线图的定义，将原数据处理成为符合 GCN 要求的数据。

4.1.1 线图转换

在前面介绍了线图（line-graph）的基本定义，这里就不再赘述。在本文的数据上，线图转换的所做的工作是将路网上的道路与道路交点进行转换，生成得到新的拓扑图，我们称其为路邻接拓扑图。路邻接拓扑图中的节点为路网中的道路，而路邻接拓扑图中的连边则代表着道路之间的连接关系。其过程如图 4-1 所示。

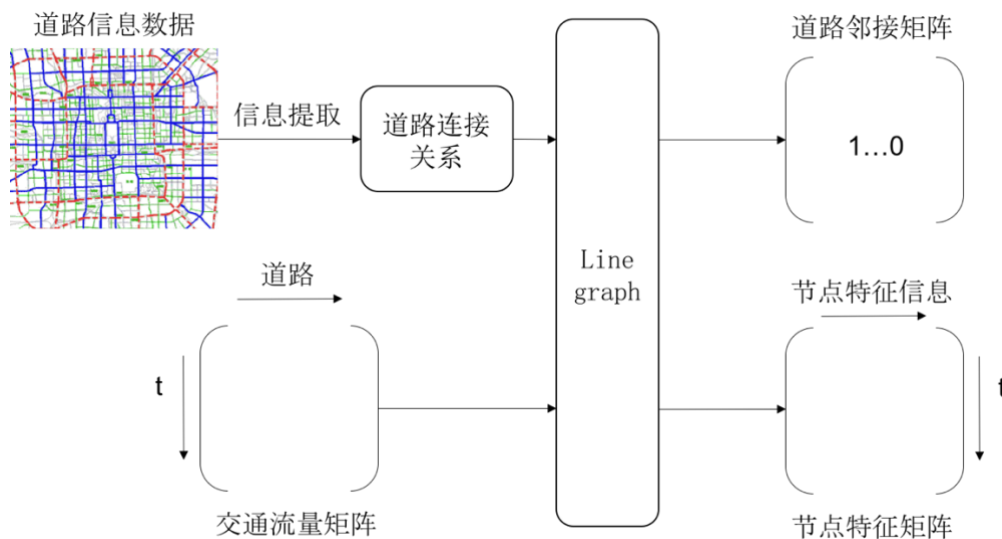


图 4-1 线图转换实际过程

Figure 4-1 The Process of Line Graph Change

在线图转换的过程中，虽然抛弃了原有的路网结构，但是得到的新的路邻接拓扑图包含了原路网结构中的所有信息。虽然路网空间信息没有改变，道路上交通流量统计得到的道路交通流量矩阵与路邻接拓扑图中的道路交通流量矩阵也没有本质上的区别，但是道路交通流量矩阵所描述的对象改变了。在原始的路网拓扑结构中，道路交通流量矩阵是作为连边的一个特征矩阵存在的。而在线图转换后得到的路邻接拓扑图中，道路交通流量矩阵是作为节点的一个特征矩阵存在的。因此进行线图转换操作的意义就是在于将道路交通流量矩阵所描述的对象从

拓扑结构中的连边转变为新的拓扑结构中的节点。

线图转换在本文中的利用解决了数据不符合输入要求的问题,为合理的提取道路交通流量数据的空间信息并以此预测道路上的交通流量提供了技术基础。

4.1.2 空间特征提取及预测

在本文中选取了 GCN 用来提取交通流量数据的空间特征,而 GCN 的训练是需要有两个输入的:一个是拓扑图节点的邻接矩阵,一个是节点的特征矩阵。在经过上一小节的线图转换操作后,就得到了路邻接拓扑图,可以从中提取出道路邻接拓扑矩阵,作为输入的拓扑图节点的邻接矩阵;同时在线图转换操作后得到了道路交通流量矩阵,将其作为输入的节点的特征矩阵。具体过程如图 4-2 所示。

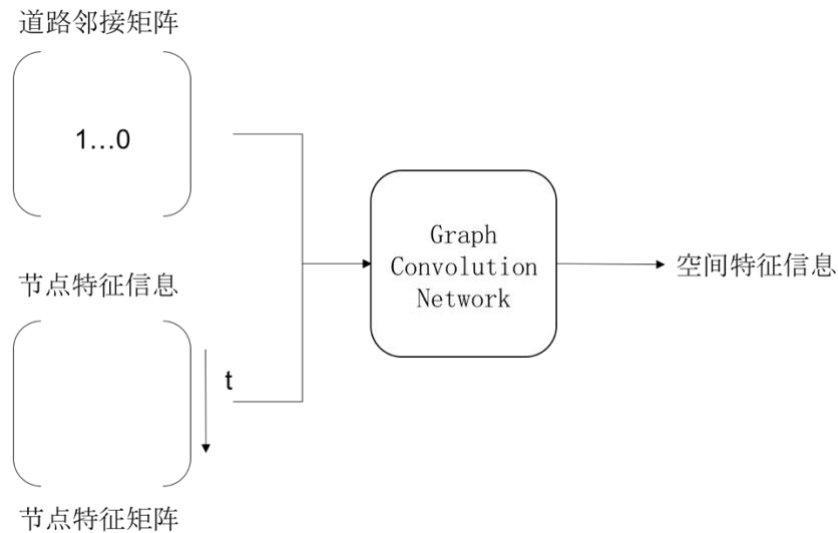


图 4-2 空间特征提取过程

Figure 4-2 The Process of Spatial Feature Extraction

在 GCN 中,需要调节的参数除了传统的神经网络中需要调节的参数,如批次大小、学习率等,最主要的需要注意的是邻居跳数的选择。正如 CNN 中用卷积核对图片的扫描来提取图片的特征,通过控制卷积核的大小来提取不同的特征,GCN 中也是利用了一个卷积的概念。在 GCN 中的表现形式就是选取某一中心点之后,控制该中心点到其他节点的跳数来控制提取不同维度的空间特征。理论上来说,GCN 所选取的跳数越大,所汲取的空尽特征信息就越多,结果应该就越好。可是在实验中发现,在跳数增加过成功中,预测结果会先变好再变差,分析产生这种结果的原因是真实的道路上其他道路对所选中心道路的影响关系随着跳数的增加而减小,还有可能引入噪声,所以跳数不能无限制的增加,在实验中

选择跳数为 3 跳，这样能得到较好的预测结果。

4.2 交通流量时间特征提取及预测

在提取时间特征方面，本文中采用了常用的 LSTM 网络，利用 LSTM 对已经提取了空间信息的序列进行处理，提取其中的时间特征。LSTM 通常被用于时间序列信息的提取，因为它通过对时间序列数据信息的提取，并合理的整合前方远处的时间序列数据信息来对下一时刻的数据进行预测，同时还克服了普通 RNN 循环神经网络的梯度消失或者梯度爆炸问题。具体过程如图 4-3 所示。

在这里使用了单层的 LSTM 网络。设置批次大小（batch_size）为 60，时间步长（time_step）为 10，即训练批次的大小是 60 个单位，每 60 个时间步就更新一次结果，而在每次训练时，是输入 10 时间步的道路流量信息，这些道路流量信息已提取了其中的空间信息，利用 LSTM 网络来预测第 11 个时间步的交通流量信息。

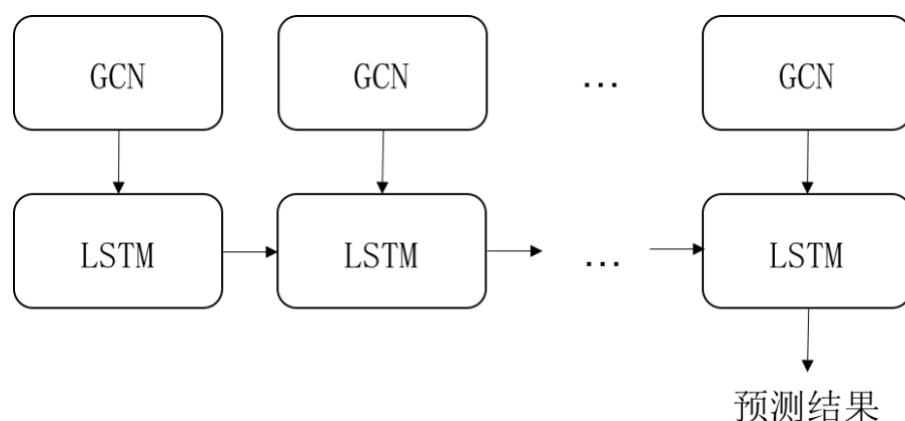


图 4-3 时间特征提取过程

Figure 4-3 The Process of Time Feature Extraction

4.3 交通流量其他特征提取及预测

根据 3.1 节中对交通流量预测的假设以及 3.2 节、3.3 节中对所拥有数据的挖掘、分析及处理，得出交通流量数据不仅仅与时间信息和空间信息有关，还有一些其他的特征，如周期特征等。虽然周期特征也属于时间特征的一种，但是在 4.2 节中提取的时间信息是短时间内的时间特征信息，而周期信息是一个长时间维度上的时间特征信息，并且由于数据分析与预处理的结果，发现数据的时间周期性是一个比较强的特征，因此决定将数据的周期性独立出来，单独提取交通流量数据的周期特征并对流量进行预测。除周期特征外，根据实际的路况信息，本

文还认为道路上的交通流量还与当天的天气有关,因此本文还将天气作为其他特征进行提取并对流量进行预测,共同作用于最终道路交通流量的预测。

4.3.1 周期特征提取及预测

在提取道路上交通流量的周期信息时,利用 Python 中的 statsmodels 工具包,其中含有针对周期性分解的函数 `seasonal_decompose`,可以将道路上的交通流量序列进行分解。

首先在代码首部输入“`from statsmodels.tsa.seasonal import seasonal_decompose`”,从 statsmodels 中调用 `seasonal_decompose` 函数。接下来利用 `decomposition = seasonal_decompose(ts, freq=144, two_sided=False)` 对道路上交通流量中的各类信息进行分解。其中 `ts` 为输入的时间序列, `freq` 为时间序列的周期,在这里本文选定周期为一天,即含有 144 个数据点。将结果做成图展示出来,如图 4-4 所示。

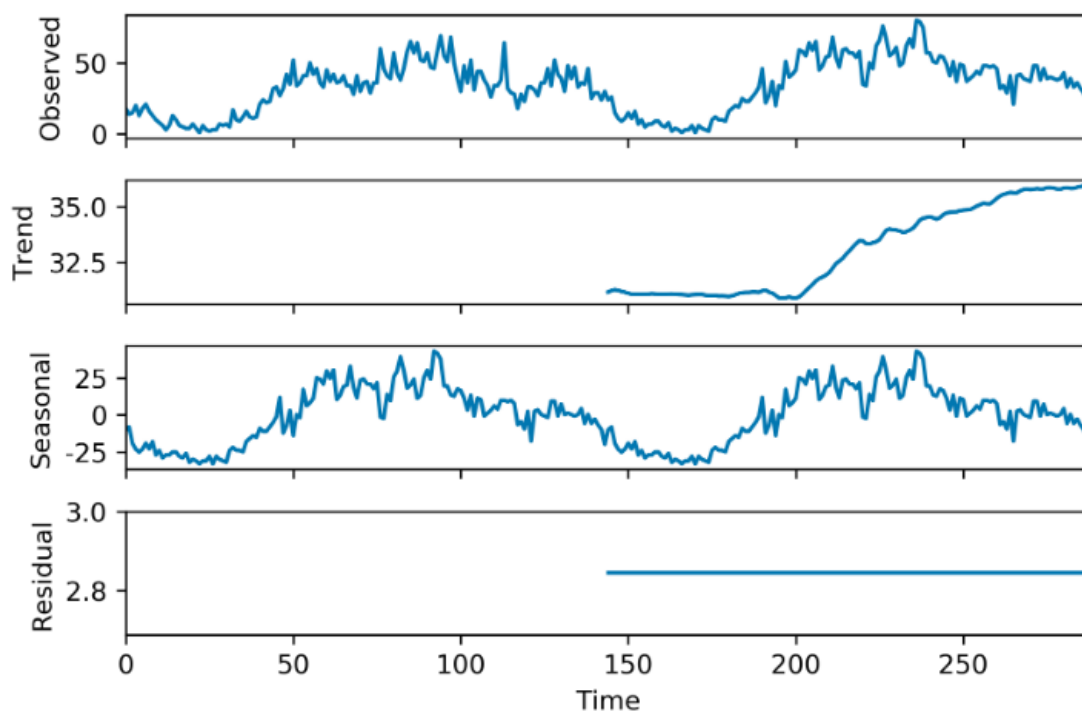


图 4-4 交通流量数据特征分解

Figure 4-4 Traffic Flow Feature Decomposition

在图 4-4 中, `observed` 是对原始时间序列的描述, `trend` 是对从时间序列分解出来的趋势部分的描述, `seasonal` 是对从时间序列分解出来的周期部分的描述, `residual` 是对从时间序列分解出来的残差部分的描述。要提取的就是 `seasonal` 所描述的周期信息特征。将周期信息特征单独提取出来,并以周期信息序列作为预

测值对第三天道路上的交通流量值进行预测,同时与第三天道路上真实的交通流量值进行比较,最终将其作为对时空特征提取预测结果的矫正加入最终的预测结果。

4.3.2 其他特征提取及预测

最后对上述特征以外的其他特征进行提取,在这里选取当天的天气信息作为本文所要提取的其他信息。实验中所应用的天气信息包含每天每个小时的天气温度,将其作为特征一同输入进 LSTM,相当于对输入 LSTM 的数据增加一维特征,来对道路上的交通流量进行预测。该结果同样作为对时空特征提取预测结果的矫正加入最终的预测结果。

4.4 预测结果整合

由 4.1-4.3 三小节得到了一条道路上交通流量的三个预测结果,分别是:基于道路上交通流量时空特征提取后的预测结果,基于道路上交通流量时间周期特征提取后的预测结果,基于道路上交通流量天气特征提取后的预测结果。接下来要对这三个结果进行整合,得到最终的道路上交通流量的预测结果。在本文的实验中,选择线性相加并计算平均值的方式得到了最终的道路上交通流量的预测结果,并以此结果作为本文所设计的基于深度学习的交通流量预测系统的输出。

4.5 本章小结

本章主要对本文所提出的交通流量预测系统进行的细致的描述,主要从特征提取的方向对交通流量预测模型进行划分,分为了流量空间特征提取及预测、流量时间特征提取及预测、流量其他特征提取及预测三部分,描述了这三部分的核心算法,并介绍这三部分中所使用的模型及参数,阐明了各部分的输入与工作机理,并对其应用的原因进行分析。最后对预测结果进行整合,最终得到本文所设计的流量预测系统的输出结果。

5 实验对比与结果分析

本章主要对本文所设计提出的交通流量预测系统进行试验评估,利用选取的评价指标,对本文设计的系统的效果进行测评。接下来选取经典的和有可比性的其他交通流量预测模型在所选用数据集上进行试验,将这些模型的结果与本文所设计提出的交通流量预测系统的结果进行对比,以此来分析不同模型在进行交通流量预测产生的该结果的原因,并结合模型算法的本质加以分析,最终阐述本文提出的交通流量预测系统的优点。同时对本文所设计提出的交通流量预测系统进行剖析,利用控制变量法分析系统内各模块对交通流量预测系统结果的影响,并对结果加以解释。最后对本章内容进行总结概括。

5.1 参数设置与结果展示

将本文所选取的数据集与设计的交通流量预测系统进行结合,利用交通流量预测系统对现有的数据进行预测。在本节中主要对系统模型的参数进行分析,并根据结果的表现选取相对最优的参数。之后将本文所设计的交通流量预测系统的预测结果进行展示,并对结果进行简要分析。

5.1.1 系统参数

在本文实验模型的超参数选取上面,主要注意了模型的学习率、隐藏层神经元个数等。其中,主要对隐藏神经元个数进行了具体剖析。隐藏层神经元个数的多少会对交通流量预测的准确率有所影响。因此,尝试使用不同的隐藏层神经元个数进行试验,利用对交通流量预测的准确率的评估结果作为评价指标,由实验得出在不同的隐藏层神经元个数情况下的交通流量预测的准确率,以此结果作为选择隐藏神经元个数的标准,选择结果最好的情况下所采用的隐藏神经元个数。在实验中,尝试在隐藏层神经元个数设置为 20、40、60、80 的情况下对道路上的交通流量进行预测,并输出交通流量预测的准确率,结果如图 5-1 所示。根据实验结果可以得出,当隐藏层神经元个数为 60 时,均方误差值最低,因此在实验中使用隐藏层神经元个数为 60 作为最终实验的超参数。

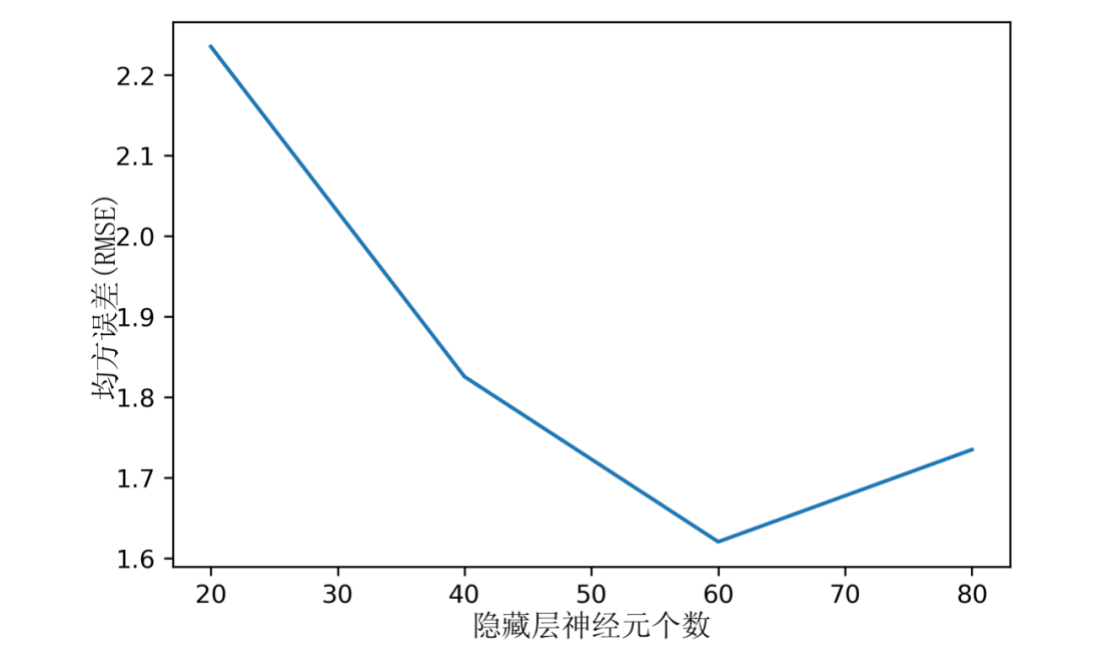


图 5-1 隐藏层神经元个数对预测结果的影响
Figure 5-1 Influence of Hidden Layer Neurons Number

设置学习率为 0.001，训练批次的大小为 60，选取 Adam 作为优化器，并选取 2/3 的数据（即两天）作为训练集输入模型对模型进行训练，其余数据作为测试集。其中，选取 Adam 优化器是因为其实现简单、计算高效、效果优异。

5.1.2 实验结果展示

将本文所设计的交通流量预测系统的性能结果用均方误差（RMSE）来进行衡量，均方误差（RMSE）的数学表达式如下：

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}$$
(5-1)

因此，最后得出各种本文所设计的交通流量预测系统的性能结果如表 5-1 所示：

表 5-1 交通流量预测系统结果评价
Table 5-1 Evaluation of Traffic Flow Prediction System

模型名称	评价指标 (RMSE)
GCN+LSTM+P+W	1.6203

根据表 5-1 中的数据结果，可以看出本文所提出设计的交通流量预测系统的预测结果与真实结果差别很小，以此可以证明本文所设计的交通流量预测系统性能较好并可以良好的应用在实际数据上，交通流量预测值与真实值的对比结果如图 5-2 所示。

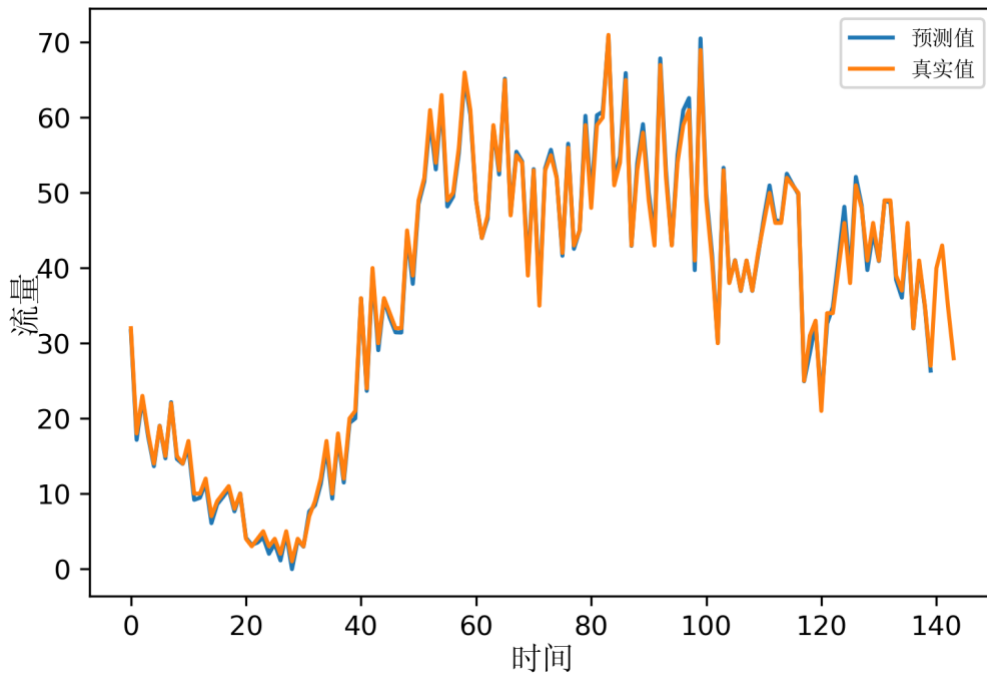


图 5-2 交通流量预测系统预测结果

Figure 5-2 Results of Traffic Flow Prediction System

5.2 模型结果对比与分析

本节主要对所选取的常用的以及有对比性的以往的交通流量预测方法作为基线模型进行实验，对其实验过程及结果进行描述，并将实验结果与本文所提出的交通流量预测系统的结果进行对比，分析本文所提出的交通流量预测系统的优点。

5.2.1 交通流量预测基线模型

除了本文所设计实现的交通流量预测系统，本文还在所拥有的数据集上实现了几种以往的常用的交通流量预测方法作为基线模型进行对比，如 HA、ARIMA、HMM、LSTM、CNN+LSTM 等。本文还提出了一种手动提取周边空间信息的 mutil-LSTM 算法，与上述以往的交通流量预测方法一起作为本文的基线对比模型。其中，HA 是历史平均值，ARIMA 基于统计学基础，HMM 基于机器学习与马尔可夫方法，LSTM、CNN+LSTM、mutil-LSTM 三种方法基于深度神经网络。三种非神经网络的方法都是只着眼与时间信息上，而基于深度神经网络的三种方法中，LSTM 模型也只是基于时间特征信息上，CNN+LSTM 和 mutil-LSTM 两种方法都是基于时间特征信息与空间特征信息的，只是在空间特征信息提取的方法上有所不同。接下来本节将对上述的几种的常用的交通流量预测算法进行描述，

并展示出这些算法在本文数据集上的效果。

(1) HA

HA (History Average) 模型就是使用数据历史时期中的平均交通流量信息作为下一时刻交通流量信息的预测方法。在 HA 模型中, 总共使用了连续三天的某一特定道路上的交通流量序列, 选去其中前两天的交通流量数据, 计算数据的平均值, 并将该平均值作为第三天交通流量数据的预测值, 与第三天的真实的交通流量数据进行对比。经计算前两天交通流量数据的平均值为 33.513888, 与第三天数据进行比较, 均方误差 (RMSE) 为 18.4332。

(2) ARIMA

在 ARIMA 模型中, 也同样使用了三天的某一特定道路上的交通流量序列, 其中利用前两天的数据去设定模型与调整, 进而来预测第三天道路上的交通流量数据并与真实的第三天的交通流量数据进行对比。

首先要检测时间序列的平稳性, 进行 ADF 检验, ADF 检验可以得到单位根检验统计量对应的 pvalue 值, 若此值显著大于 0.05, 则该序列非平稳。并分析序列中的白噪声信息, 做出自相关系数图与偏自相关系数图, 结果如图 5-3、图 5-4。并根据图 5-3、图 5-4 的结果来确定模型的 p、q 值。

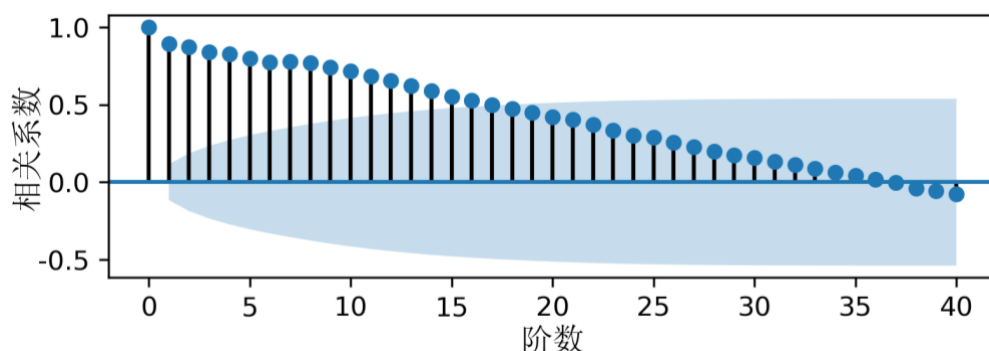


图 5-3 自相关系数图

Figure 5-3 Autocorrelation Coefficient

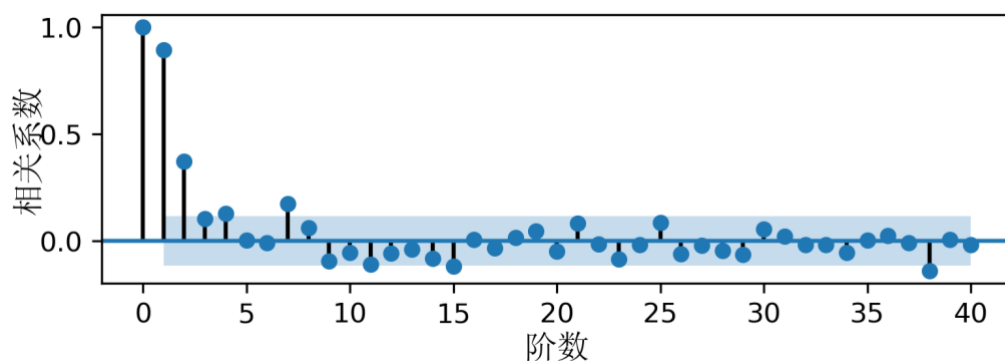


图 5-4 偏自相关系数图

Figure 5-4 Autocorrelation Coefficient

之后进行差分阶数 d 选择,若差分后的序列不平稳,修改差分阶数重新差分,或者继续对序列进行差分运算,运算后继续进行平稳性和白噪声检验,直至差分后的序列平稳为止;若差分运算后的序列平稳,则对差分后的平稳序列拟合 ARMA 模型,并且为模型定阶。

在实验中, p 值选择范围 7、8, q 值选择范围 0、1, d 值选择范围 0、1,经过实验对比各组模型的 AIC、BIC、HQIC 值,最终选择模型 ARIMA (7, 0, 0) 作为实验数据的预测模型。最后得出的预测结果如图 5-5 所示,均方误差(RMSE)值为 9.5155。

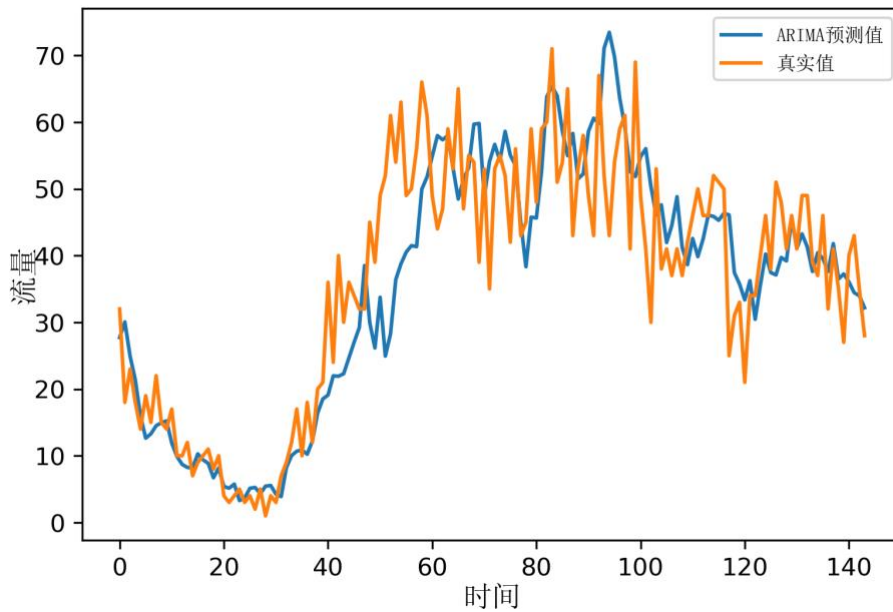


图 5-5 ARIMA 模型预测结果

Figure 5-5 Prediction of ARIMA

(3) HMM

在 HMM 模型中,同样利用前两天的数据对模型进行训练,进而预测第三天的交通流量数据,并用第三天真实的道路上的交通流量值作为对照,计算 HMM 对道路上交通流量预测的准确率。

在本文使用的 HMM 模型中,设置了 10 个隐藏状态,并设置相邻时间点流量的对数差值至每间隔 6 个时间点流量的对数差值等 6 个特征。将各个特征的长度进行统一,并将统一长度后的各个特征进行拼接。这之后利用 `model=hmm.GaussianHMM().fit()` 对模型进行训练,其中隐藏状态数 `n_components=10`,协方差类型 `covariance_type="full"` (是指在每个马尔可夫隐含状态下,可观察态向量使用完全协方差矩阵,对应的协方差矩阵里面的元素都是不为零。`full` 是最强大的,但是需要足够多的数据来做合理的参数估计),要执行的最大迭代次数 `n_iter=20000`。模型训练之后利用 `hidden_states=model.predict()`

输出该模型的隐藏状态，如图 5-6 所示。接下来输出在训练数据中各个数据点在真实数据上各个状态的分布，如图 5-7 所示。

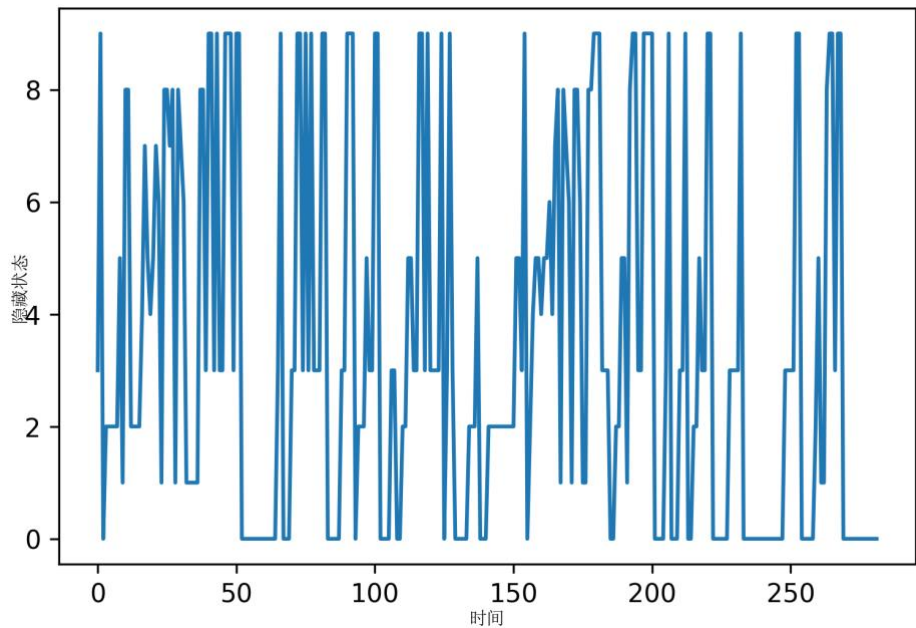


图 5-6 时间序列中各个值的隐藏状态
Figure 5-6 Hidden States of Values in Time Series

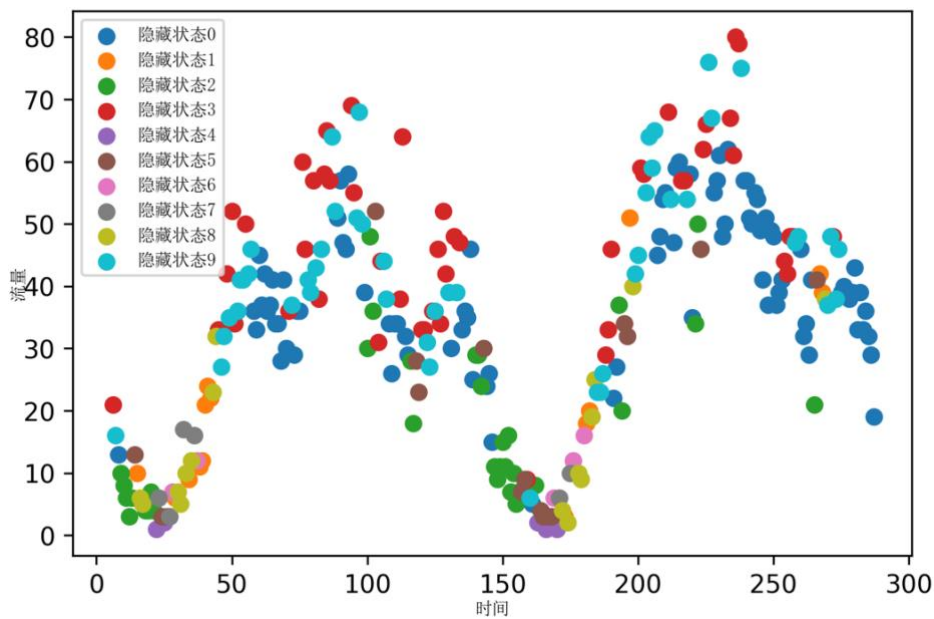


图 5-7 时间序列中隐藏状态分布
Figure 5-7 Distribution of Hidden States in Time Series

接下来要将各个状态的特征打印出来以此来选择该隐藏特征所负责的内容是增加还是减少。其中由于 4、8 状态变化十分明显表现为上升，因此只打印出剩余状态的变化图像，如图 5-8 所示。

根据图 5-8 及前面分析可知，状态 0、1、4、8、9 这五个状态表现出上升趋势

势，状态 2、3、5、6、7 这五个状态表现出下降趋势。因此设定当出现状态 0、1、4、8、9 这五个状态时，让流量值呈现上升状态；当出现状态 2、3、5、6、7 这五个状态时，让流量值呈现下降状态。接下来设置验证曲线，若所选择的趋势是正确的，生成的验证曲线应该是不断上升的。验证曲线如图 5-9 所示。根据图 5-9 可知，所划分的状态趋势是正确的。

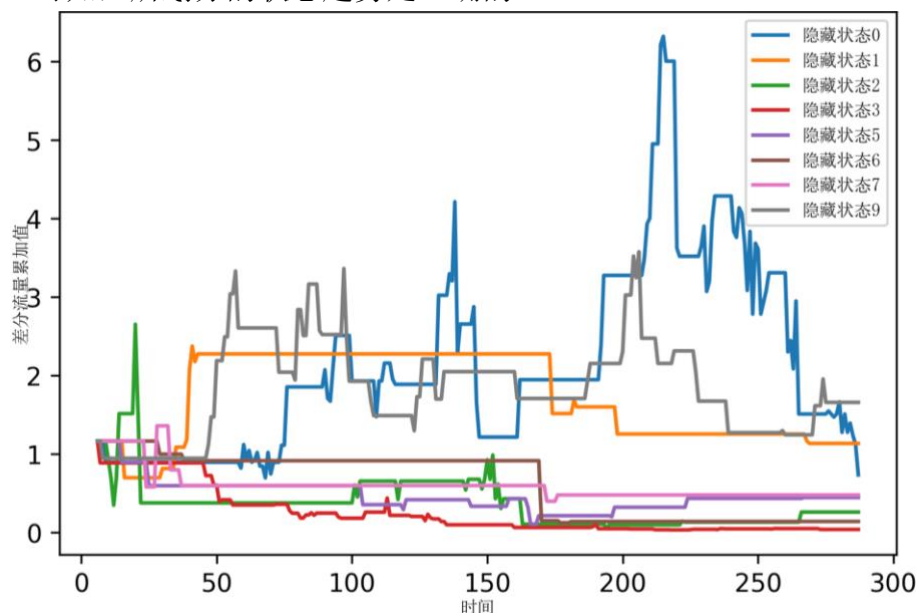


图 5-8 各隐藏状态变化趋势

Figure 5-8 Trend of Hidden States Changing

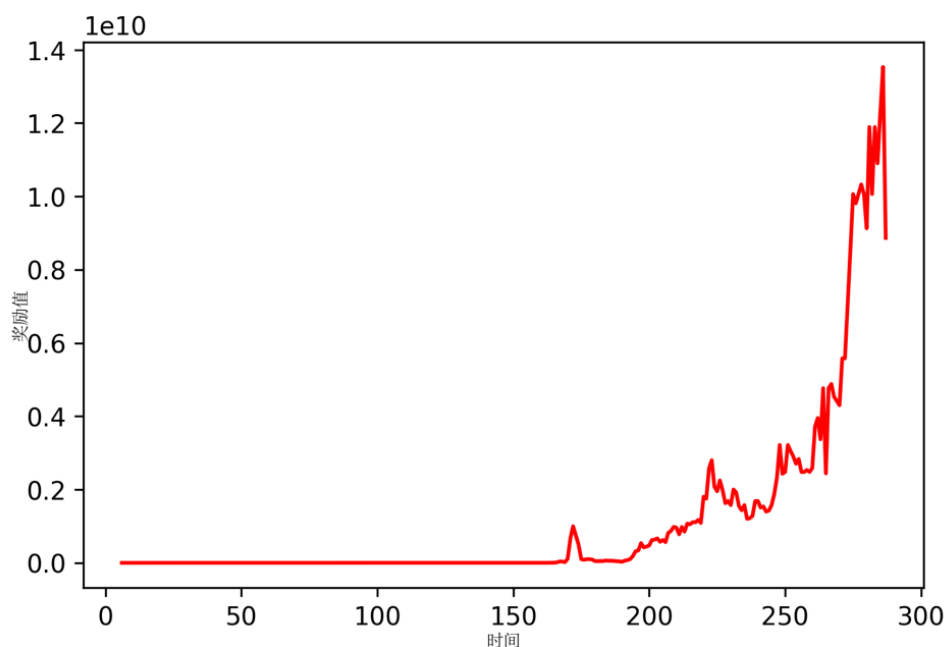


图 5-9 隐藏状态设置结果验证

Figure 5-9 Verification of Hidden States Setting

在此之后利用 `model.sample()` 函数产生预测值，其中预测数量 `n_samples=144`，即第三天数据点的数量。最终将预测结果与真实结果花在同一张图中，结果如图

5-10 所示，均方误差（RMSE）为 27.01286480。

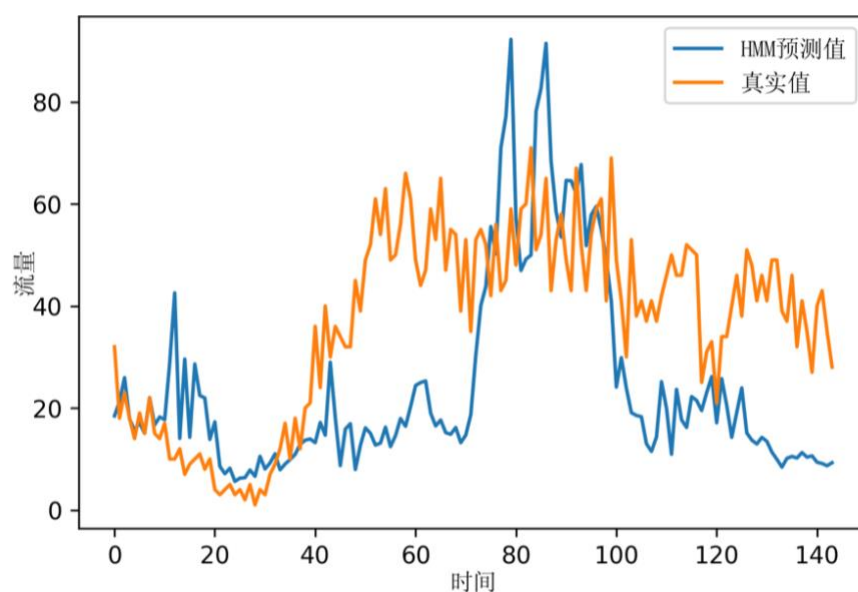


图 5-10 HMM 模型预测结果

Figure 5-10 Prediction of HMM

(4) LSTM

在 LSTM 模型中，同样也是利用前两天的数据对模型进行训练，对第三天的交通流量数据进行预测，并与第三天真实的交通流量数据进行对比。设置 `batch_size=60`，`time_step=10`，即利用前十个时间步的数据来预测第十一个时间步的数据。LSTM 网络中利用均方误差作为训练时的 loss 值，利用 Adam 作为优化器，利用 RMSE 作为模型预测准确度的衡量标准。结果如图 5-11 所示，均方误差（RMSE）为 6.7031。

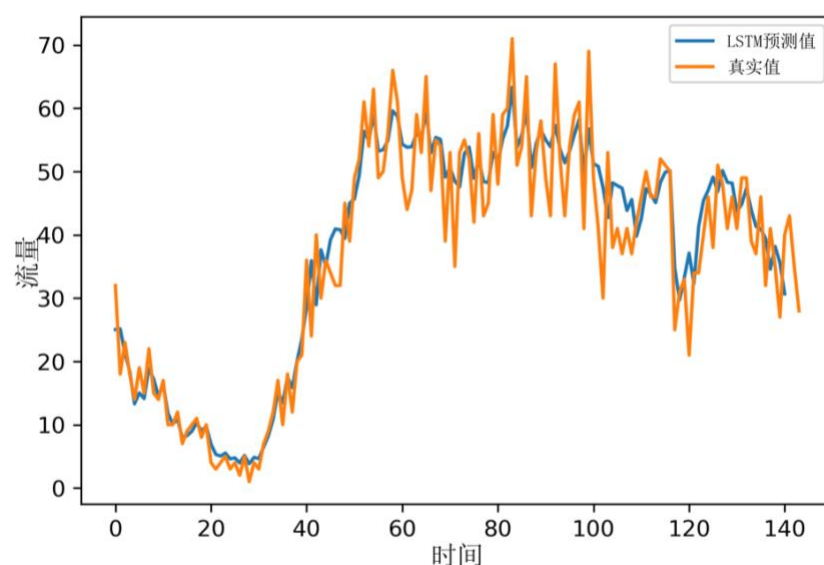


图 5-11 LSTM 模型预测结果

Figure 5-11 Prediction of LSTM

(5) CNN+LSTM

在 CNN+LSTM 模型中, 将车辆信息数据利用经纬度信息与地图进行结合上, 之后利用经纬度信息对车辆信息数据进行切分, 统计切分后得到的小块中车辆的流量值作为该块中的特征值。之后使用 CNN 来提取其中的空间特征, 再利用 LSTM 提取其中的时间特征。。

在实验过程中, 将数据处理得到符合输入的元数据后, 先对数据进行最大最小归一化, 使得流量值归一化后都落在区间 $[0,1]$ 之间, 避免数据较为分散, 跨度过大, 同时提升模型的收敛速度与精度。之后设置三层 CNN, 初始层维度为 32×32 即 1024 维, 滤波器数量为 128, 卷积核大小为 3×3 , 激活函数为 ReLu()函数。经过三层卷积层之后是一个最大池化层, 其中池化大小为 4。接下来就是接入 LSTM, 其中激活函数为 ReLu 函数。最后将 LSTM 的输出结果介入一个 Dense 层, 对特征进行整合。CNN+LSTM 模型预测结果如图 5-12 所示, 其均方误差 (RMSE) 为 4.3512。

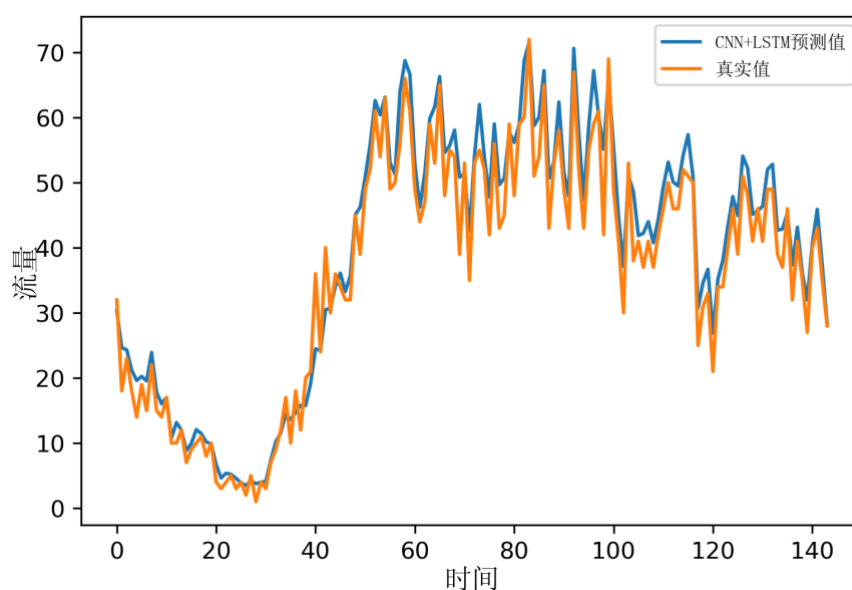


图 5-12 CNN+LSTM 模型预测结果

Figure 5-12 Prediction of CNN+LSTM

(6) mutil-LSTM

mutil-LSTM 模型是本文自己设计的模型, 通过手工提取中心道路以及与中心道路连接道路的道路交通流量值一起作为特征值送入 LSTM 网络中, 这相当于手工提取了中心道路及其周边一阶邻接道路的作为空间特征, 然后利用 LSTM 来提取其中的时间特征。这与本文所提出的道路交通流量预测方向是有相似的地方的, 因为手工提取的空间信息是真实的道路空间信息, 与本文中的方法提取的道路交通流量空间信息的一跳邻接空间信息几乎完全一致, 只是本文提取

的道路交通流量空间信息包含的信息更多。因此 mutil-LSTM 模型这种方法拥有很强的对比性，所以在这里引入这种模型与本文的模型做对比。

在 mutil-LSTM 模型的参数设置方面，选择与上述的 LSTM 模型相同的神经元个数、time_step、batch_size、学习率等超参数，只是将每一时间步的输入由一个值变为五个值，同样利用前两天的数据对模型进行训练，来对第三天的数据进行预测。mutil-LSTM 的预测结果如图 5-13 所示，其均方误差(RMSE)为 2.3669。

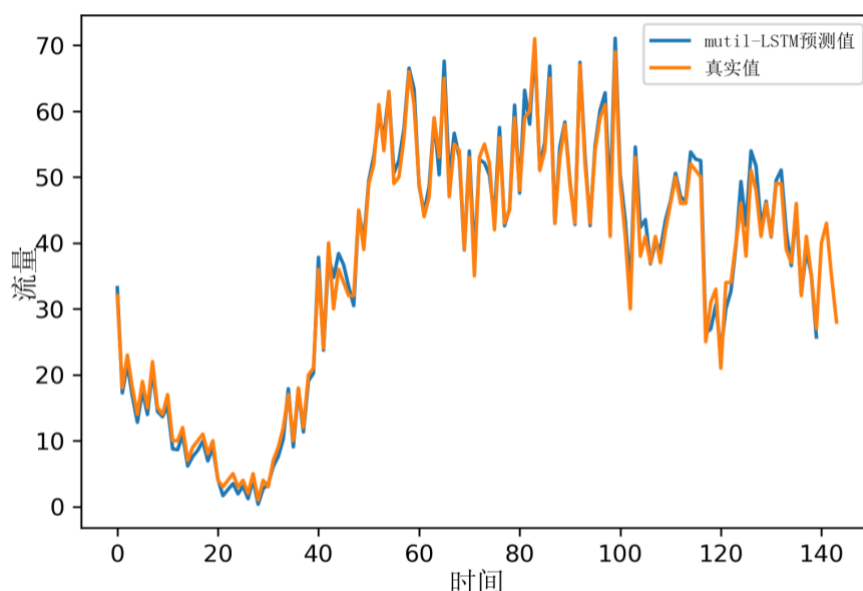


图 5-13 mutil-LSTM 模型预测结果

Figure 5-13 Prediction of mutil-LSTM

5.2.2 结果对比分析

通过以上对本文提出的交通流量预测系统与各基线交通流量预测方法的介绍与实验，最终将本文所涉及到的相关的道路上交通流量预测的模型与实验结果进行整合，得到本文所设计的交通流量预测系统与各种交通流量预测基线方法的性能结果如表 5-2 所示。

根据表 5-2 中的数据结果，可以看出本文所提出设计的交通流量预测系统相对于以往的交通流量预测方法来说表现效果最优，均方误差值低于上述所有的基线交通流量预测方法。对该结果进行深入研究，可以得出以下结论：

(1) 预测效果明显优于其他模型。通过对比本文提出的交通流量预测系统与以往的交通流量预测模型的 RMSE，可以看出本文所设计的交通流量预测系统效果优于以往的模型。产生这种结果的原因是本文所设计的交通流量预测系统合理的提取了道路上交通流量的空间特征，并结合了道路上交通流量的时间特征进行预测，同时结合道路上交通流量的其他特征的预测结果，共同作用于下一时刻道

路上的交通流量。

(2) 时空特征效果明显。在空间特征方面，为了验证空间特征是否处于不可或缺的地位，将本文所提出的交通流量预测系统和 CNN+LSTM、mutil-LSTM、GCN+LSTM 共四个模型的结果作为一组整体与其他模型的结果进行对比。通过结果对比发现，提取空间特征用来预测道路上的交通流量是十分有效的；而通过 CNN+LSTM 模型的预测效果与本文所提出的交通流量预测系统的预测效果进行对比，可以说明本文所提出的交通流量预测系统中的空间特征提取方法相比于 CNN 来说能够更好的提取道路交通流量的空间特征；通 mutil-LSTM 模型的预测效果与本文所提出的交通流量预测系统的预测效果进行对比，可以说明本文所提出的交通流量预测系统能够整合多阶邻接道路的空间特征，以实现道路上的交通流量更好的预测。在时间特征方面，通过对 HA 模型与剩余其他模型的结果进行对比分析，表明提取时间特征用来预测道路上的交通流量是十分有效的。通过对 HA、ARIMA、HMM 模型与剩余其他模型的结果进行对比分析，可以得出利用了神经网络的模型的效果相比其他模型而言对道路上交通流量进行预测的效果更好，这主要是由于 HA，ARIMA 和 HMM 等方法难以处理复杂的非平稳时间序列数据。

表 5-2 各模型交通流量预测结果评价
Table 5-2 Evaluation of Traffic Flow Prediction of Various Models

模型名称	评价指标 (RMSE)
HA	18.4332
ARIMA	9.5155
HMM	27.0129
LSTM	6.7031
CNN+LSTM	4.3512
Mutil-LSTM	2.3669
GCN+LSTM	1.8774
GCN+LSTM+P+W	1.6203

综上，对上述结果的解释是本文所提出的交通流量预测系统相对于以往的交通流量预测模型而言对预测道路上的交通流量更为准确，因为本文所提出的交通流量预测系统能够更好的提取道路交通流量的空间特征并结合道路交通流量的时间特征进行预测，并结合利用时间周期特征以及天气信息特征等其他特征等预测结果共同作用得到最终的道路上交通流量的预测值。

5.3 各模块结果与功能分析

上一小节对本文所提出设计的交通流量预测系统与以往的交通流量预测方法的效果进行对比，详细的说明了本文所提出设计的交通流量预测系统的优点。本节主要是对本文所提出设计的交通流量预测系统内部进行剖析，分析系统内本文所提出的空间特征提取模块与其他特征提取模块的功能及作用并加以解释，并分别分析其他特征提取模块中时间周期特征提取模块与天气信息特征提取模块的性能与作用，以此来证明本文所提出设计的交通流量预测系统中各个模块都是不可或缺的。

5.3.1 空间特征提取模块

在本文所提出设计的交通流量预测系统中，空间特征提取模块主要由图卷积网络（GCN）构成，该部分主要是用以合理的提取道路上交通流量信息的空间特征。在实验中，通过是否接入空间特征提取模块（即 GCN 网络）来验证确定空间特征提取模块的作用。表 5-3 显示出包含空间特征提取模块的交通流量预测系统与不包含空间特征提取模块的交通流量预测系统的结果对比，不包含空间特征提取模块的交通流量预测系统的流量预测均方误差为 6.5362。从此对比结果中可以分析出，本文设计的空间特征提取模块在提升交通流量预测结果方面有显著的效果。

表 5-3 空间特征提取模块性能分析
Table 5-3 Performance Analysis of Spatial Feature Extraction Module

模型名称	评价指标（RMSE）
GCN+LSTM+P+W	1.6203
LSTM+P+W	6.5362

5.3.2 其他特征提取模块

在本文所提出设计的交通流量预测系统中，其他特征提取模块主要由周提特征提取模块与天气信息特征提取模块两部分共同组成。该模块主要是用来提取一些与道路上交通流量有关其他的辅助信息，来辅助道路交通流量预测，提高预测准确率。通过是否接入其他特征提取模块（即周提特征提取模块与天气信息特征提取模块）来确定其他特征提取模块的作用。表 5-4 显示出包含其他特征提取模块的交通流量预测系统与不包含其他特征提取模块的交通流量预测系统的结果

对比，不包含其他特征提取模块的交通流量预测系统的流量预测均方误差为1.8774，从此对比结果中可以分析出，本文设计的其他特征提取模块在交通流量预测方面有显著的效果。

表 5-4 其他特征提取模块性能分析
Table 5-4 Performance Analysis of Other Feature Extraction Module

模型名称	评价指标 (RMSE)
GCN+LSTM+P+W	1.6203
GCN+LSTM	1.8774

在其他特征提取模块中，再分析其中的周提特征提取模块与天气信息特征提取模块这个两个模块分别对道路交通流量预测准确率提升所做的贡献。表 5-5 显示出包含其他特征提取模块的交通流量预测系统与不包含其他特征提取模块中的周提特征提取模块和不包含其他特征提取模块中的天气信息特征提取模块的交通流量预测系统的结果对比，不包含其他特征提取模块中的周提特征提取模块的交通流量预测系统的流量预测均方误差为 1.6376，不包含其他特征提取模块中的天气信息特征提取模块的交通流量预测系统的流量预测均方误差为 1.8594。从此对比结果中可以分析出，本文设计的其他特征提取模块中的周提特征提取模块在提升交通流量预测的准确率方面有着显著的效果，而其他特征提取模块中的天气信息特征提取模块在提升交通流量预测的准确率方面效果不明显。产生这种结果的原因是周期信息是着重提取了道路交通流量时间信息中的周期信息，而且通过数据观察后发现道路上的真实交通流量是呈周期性的，说明周期特征是道路上的真实交通流量的一个强特征。而天气信息虽然根据我们的预估认为会对交通流量的预测结果也会产生一些影响，但是在本文的实验过程中，所使用的数据所覆盖的时间上天气信息并没有较大的变化，因此并不能十分完美的体现天气信息特征在道路交通流量预测上所产生的作用，所以才会出现周期信息特征对提升交通流量预测的准确率有较大的作用而天气信息特征对提升交通流量预测的准确率作用较小的情况。

表 5-5 周期特征提取模块与天气特征提取模块性能分析
Table 5-5 Performance Analysis of Periodic and Weather Feature Extraction Module

模型名称	评价指标 (RMSE)
GCN+LSTM+P+W	1.6203
GCN+LSTM+P	1.6376
GCN+LSTM+W	1.8594

5.4 本章小结

本章主要对本文所涉及的实验进行归纳总结并进行对比,同时对产生结果的原因进行分析。首先说明了本文所设计的交通流量预测系统中所涉及的参数,并对参数进行试验分析,选取最为合适的参数用于最终的实验,并对本文所设计的交通流量预测系统的结果进行展示分析。接下来对本文选取以往的交通流量预测模型进行分析及实验结果的展示,并与本文所设计的交通流量预测系统的实验结果进行比较分析,得出本文所设计的交通流量预测系统效果最佳并分析得到相应的结论。然后对本文所设计的交通流量预测系统中空间特征提取模块与其他特征提取模块的作用并进行分析,利用控制变量法实现输出缺失某一模块时该模型对道路上的交通流量预测的准确性,进而分析该模块在系统中所产生的作用。在分析其他特征提取模块的功能过程中,对其他特征提取模块进行解析,分析其中的周期特征提取模块与天气信息特征提取模块分别对道路交通流量预测准确率提升所做的贡献,并对结果进行合理的解释。

6 总结及展望

本章主要对本文所做的工作进行了总结归纳并对本文相关方向的未来发展进行了展望。主要是对本文的贡献进行了归纳总结，详细的说明了本文所做的工作并列出了的本文所达成的关键性成果。同时也分析了本文实验中所存在的问题，主要包含尚待改进的地方和尚未解决的问题，并以此结合文本所研究内容的发展方向进行展望，同时对本文所提出的交通流量预测系统还可以改进的地方进行了阐述。

6.1 论文工作总结

本节主要对本文所涉及的相关工作进行总结并对本文达成的成果进行了归纳。

随着大数据、云计算、人工智能等信息技术的快速发展，为解决城市交通拥堵问题，智能交通系统的规划建设也进入高速发展时期，将是未来出行和生活方式的发展方向。而流量预测系统正是基于大数据与人工智能共同作用的智能交通系统的基础系统，正因为流量预测系统的存在才使得智能交通系统得以完整的运作。面对现有的需求与条件，本文的研究目的为更好解决流量预测问题，使得我们能够更好的对道路上的交通流量进行预测。根据上述研究目的，本文设立了主要需要解决的 3 个挑战，包括：如何将车辆数据集合理的转换成道路上的交通流量信息；如何合理的提取道路上的交通流量信息的特征，包含空间特征、时间特征与其他特征；如何能够效果更好的对道路上的交通流量进行预测，以提高道路上交通流量预测的准确率。

针对上述提出的挑战，本文完成了以下的工作并做出了以下的贡献：

(1) 有效的将车辆数据集转换成道路上的交通流量并巧妙的、更切合实际的提取了道路上的交通流量的空间特征。

本文通过对车辆数据集的清洗及整形处理，将车辆数据集整形成道路上的交通流量，通过统计学观察后发现了道路上的交通流量在时间维度、空间维度以及其他维度上所表现出的特征，并分析特征提取顺序。在空间特征提取过程中，本文对道路上的交通流量的空间特征信息进行分析，并提取道路上的交通流量的空间特征。根据道路上的交通流量的实际空间属性，将路网视为拓扑图，利用线图（line-graph）的方式将拓扑图中的节点与连边进行转换，此时道路上的交通流量特征即转变成为新拓扑图中的节点信息特征，对这些信息进行统计结合就得到了

节点特征矩阵，并根据线图转换后得到的新的拓扑图生成转换后拓扑图的邻接矩阵，此邻接矩阵即表示路网中道路的邻接关系。将得到的节点特征矩阵与道路邻接关系矩阵一同输入图卷积网络，提取交通流量信息的空间特征。在车辆数据集上进行实验分析，均方误差为 1.8774，预测误差相比于 CNN+LSTM 模型的预测误差降低了 56.9%。

(2) 设计并实现了基于深度学习的交通流量预测系统。

本文将通过数据统计观察得到的特征分维度进行提取。最终确定先对道路上的交通流量信息的空间特征进行提取，再对道路上的交通流量信息的时间特征进行提取及预测，而对道路上的交通流量信息的其他特征进行并行提取及预测。结合上述对道路上的交通流量的空间特征进行提取的方法，将已经提取出道路上的交通流量的空间特征的数据输入长短期记忆（LSTM）网络对道路上的交通流量的时间特征进行提取并进行预测。同时利用道路上的交通流量的周期特征与天气信息特征等其他特征一同对道路上的交通流量进行预测。最后将这些结果结合共同作用于对道路上的交通流量的最终预测，以此来提高交通流量预测的准确性。在车辆数据集上进行实验分析，均方误差为 1.6203，相比于 CNN+LSTM 模型的预测误差降低了 62.8%，相比于（1）中提出的预测方法预测误差降低了 13.7%。

6.2 未来工作展望

本文主要在合理的提取道路上交通流量的空间特征以及提出了一种对道路上交通流量预测的系统等方面取得了一定的研究成果，虽然效果相对于以往的方法来说提升显著，但是还存在部分的问题仍需要继续进行研究。

(1) 本文在针对道路上交通流量数据的空间特征信息的提取时，将通过线图（line-graph）转换的流量特征矩阵（即节点特征矩阵）与网络邻接矩阵输入了图卷积网络（GCN）对空间特征进行提取。虽然这种方法已经能够合理的提取道路上的交通流量的空间信息并对下一时刻的道路上的交通流量进行预测，但是输入图卷积网络（GCN）的节点特征矩阵中仅包含一维特征，即道路上的交通流量值。若能加入其他可以对道路上交通流量产生影响的空间特征，如：道路信息、周边场景信息等共同作用于道路上的交通流量信息空间特征提取，将会对道路上的交通流量预测结果产生更积极的影响。

(2) 本文在设计基于深度学习的交通流量预测系统时整合了道路上交通流量的空间特征、时间特征与其他特征的预测结果。在时间特征提取时选择了 LSTM 对流量序列中的时间特征进行提取，若对时间特征提取的方法进行探究，有可能

得到更好的交通流量预测效果。同时还可以增加其他的特征提取模块来对道路上的交通流量进行预测，如工作日-节假日等。

(3) 本文所设计的基于深度学习的交通流量预测系统在进行多特征预测结果整合时，只单纯使用了线性相加，若采用机器学习的思想，对融合比例进行学习，有可能使得预测结果得到更好的改善。

参考文献

- [1] 中华人民共和国国务院. 国家中长期科学和技术发展规划纲要(上)[J]. 中国集成电路, 2006, 15(3):1-17.
- [2] Kalamaras I, Zamichos A, Salamanis A, et al. An Interactive Visual Analytics Platform for Smart Intelligent Transportation Systems Management[J]. IEEE Transactions on Intelligent Transportation Systems, 2017(99): 1-10.
- [3] Ge H X, Dai S Q, Dong L Y, et al. Stabilization Effect of Traffic Flow in an Extended Car-following Model Based on an Intelligent Transportation System Application.[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2004, 70(6):066134.
- [4] Lee W H, Tseng S S, Shieh W Y. Collaborative Real-time Traffic Information Generation and Sharing Framework for the Intelligent Transportation System[J]. Information Sciences, 2010, 180(1):62-70.
- [5] Moon J W. On the Line-Graph of the Complete Bigraph[J]. Annals of Mathematical Statistics, 1963, 34(2):664-667.
- [6] Westin K. Intelligent Transportation System and Traffic Safety - Drivers Perception and Acceptance of Electronic Speed Checkers[J]. Transportation Research Part C, 1999, 7(2-3):131-147.
- [7] 陆化普, 孙智源, 屈闻聪. 大数据及其在城市智能交通系统中的应用综述[J]. 交通运输系统工程与信息, 2015, 15(5):45-52.
- [8] 卢志义, 聂惟聪, 陈丽珍. 基于 ARMA 模型的城市轨道交通客流量预测[J]. 河南科学, 2018, v.36; No.234(05):12-17.
- [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] Kumar S V, Vanajakshi L. Short-term Traffic Flow Prediction Using Seasonal ARIMA Model with Limited Input Data[J]. European Transport Research Review, 2015, 7(3):21.
- [11] 姚智胜, 邵春福, 高永亮. 基于支持向量回归机的交通状态短时预测方法研究[J]. 北京交通大学学报, 2006, 30(3):19-22.
- [12] Qi Y , Ishak S . A Hidden Markov Model for Short Term Prediction of Traffic Conditions on Freeways[J]. Transportation Research Part C: Emerging Technologies, 2014, 43:95-111.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet Classification with Deep Convolutional Neural Networks, [J] Advances in neural information processing systems, 2012, pp. 1097-1105.
- [14] Szegedy C , Liu W , Jia Y , et al. Going Deeper with Convolutions[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015.
- [15] He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[J]. 2015.
- [16] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2014.
- [17] Girshick R. Fast R-CNN[J]. Computer Science, 2015
- [18] Ren S , He K , Girshick R , et al. Faster R-CNN: Towards Real-Time Object Detection with

- Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6):1137-1149.
- [19] Kaiming H , Georgia G , Piotr D , et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018:1-1.
- [20] Hinton G , Deng L , Yu D , et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups[J]. IEEE Signal Processing Magazine, 2012, 29(6):82-97.
- [21] Amodei D, Anubhai R, Battenberg E, et al. Deep Speech 2: End-to-End Speech Recognition in English and Mandarin[J]. Computer Science, 2015.
- [22] Xiong W , Droppo J , Huang X , et al. Achieving Human Parity in Conversational Speech Recognition[J]. IEEE/ACM Transactions on Audio Speech & Language Processing, 2016, PP(99).
- [23] Lee J Y, Dernoncourt F. Sequential Short-Text Classification with Recurrent and Convolutional Neural Networks[J]. 2016.
- [24] Wolk K, Marasek K. Neural-based Machine Translation for Medical Text Domain. Based on European Medicines Agency Leaflet Texts ☆[J]. Procedia Computer Science, 2015, 64:2-9.
- [25] Kalchbrenner N , Grefenstette E , Blunsom P . A Convolutional Neural Network for Modelling Sentences[J]. Eprint Arxiv, 2014, 1.
- [26] Sutskever I , Vinyals O , Le Q V . Sequence to Sequence Learning with Neural Networks[J]. 2014.
- [27] Y. Wu, H. Tan, L. Qin, B. Ran, Z. Jiang, A Hybrid Deep Learning Based Traffic Flow Prediction Method and its Understanding, Transp. Res. Part C Emerg. Technol. 90 (2018) 166–180.
- [28] 邓烜堃, 万良, 丁红卫,等. 基于深度学习的交通流量预测研究[J]. 计算机工程与应用.
- [29] Gang Y , Yunpeng W , Haiyang Y , et al. Short-Term Traffic State Prediction Based on the Spatiotemporal Features of Critical Road Sections[J]. Sensors, 2018, 18(7):2287-.
- [30] Tan Z, Li R. A Dynamic Model for Traffic Flow Prediction Using Improved DRN[J]. 2018.
- [31] Feng J, Li Y, Zhang C, et al. DeepMove: Predicting Human Mobility with Attentional Recurrent Networks[C]// World Wide Web Conference. 2018:1459-1468.
- [32] Alahi A, Goel K, Ramanathan V, et al. Social LSTM: Human Trajectory Prediction in Crowded Spaces[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016:961-971.
- [33] Lv Z, Xu J, Zheng K, et al. LC-RNN: A Deep Learning Model for Traffic Speed Prediction[C]// Twenty-Seventh International Joint Conference on Artificial Intelligence {ijcai. 2018:3470-3476.
- [34] Cui Z, Ke R, Wang Y. Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-wide Traffic Speed Prediction[J]. 2017.
- [35] Yu H, Wu Z, Wang S, et al. Spatiotemporal Recurrent Convolutional Networks for Traffic Prediction in Transportation Networks[J]. Sensors, 2017, 17(7).
- [36] Chen Y, Chen F, Ren Y, et al. Poster: DeepTFP: Mobile Time Series Data Analytics based Traffic Flow Prediction[C]// The, International Conference. 2017:537-539.
- [37] Zhang J , Zheng Y , Qi D . Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction[J]. 2016.
- [38] Yu B , Yin H , Zhu Z . Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting[J]. 2017.

- [39] 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.
- [40] Yuan Z , Zhou X , Yang T . Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data[C]In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 984–992
- [41] David K. Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on Graphs via Spectral Graph Theory. [C] Applied and Computational Harmonic Analysis, 30(2):129–150, 2011.
- [42] Michael Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. [C] Advances in neural information processing systems (NIPS), 2016.

作者简历及攻读硕士/博士学位期间取得的研究成果

一、作者简历

魏中锐，男，1995 年 11 月生。2013 年 9 月至 2017 年 7 月就读于北京交通大学电子信息工程学院通信工程专业，取得工学学士学位。2017 年 9 月至 2019 年 6 月就读于北京交通大学电子信息工程学院电子与通信工程专业，研究方向是信息网络，取得工学专业硕士学位。在就读专业硕士学位期间，主要从事深度学习算法与交通流量预测方面的研究工作。

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京交通大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名:  签字日期: 2019 年 5 月 28 日

学位论文数据集

表 1.1: 数据集页

关键词*	密级*	中图分类号	UDC	论文资助
深度学习；交通流量预测；图卷积网络；线图	公开			
学位授予单位名称*		学位授予单位代码*	学位类别*	学位级别*
北京交通大学		10004	工学	硕士
论文题名*		并列题名		论文语种*
基于深度学习的交通流量预测				中文
作者姓名*	魏中锐		学号*	17125063
培养单位名称*		培养单位代码*	培养单位地址	邮编
北京交通大学		10004	北京市海淀区西直门外上园村 3 号	100044
工程领域*		研究方向*	学制*	学位授予年*
电子与通信工程		信息网络	2 年	2019
论文提交日期*	2019.06.03			
导师姓名*	郭宇春		职称*	教授
评阅人	答辩委员会主席*		答辩委员会成员	
	孙强		赵永祥、李纯喜、郑宏云、张立军	
电子版论文提交格式 文本 (√) 图像 () 视频 () 音频 () 多媒体 () 其他 () 推荐格式: application/msword; application/pdf				
电子版论文出版 (发布) 者		电子版论文出版 (发布) 地		权限声明
论文总页数*	57			
共 33 项, 其中带*为必填数据, 为 21 项。				