# 绪论

## 研究背景

随着城市中的人越来越多，伴随而来的是各种问题的出现，其中一个与人们生活息息相关的问题就是交通问题。交通问题无疑是在城市中生活的有车一族们比较关注的一个问题。面对这一问题，人们有着很多不同的解决方案，目前解决交通拥堵问题的政策有限号、摇号等措施，但是这些方法只是治标不治本，随着车辆数目的不断增加，如何更有效的将交通问题解决就成了人们不断研究和探索的问题。随着大数据、深度学习的应用不断深入，人们又陆续提出其他各种基于强大计算机硬件支持的方案来更好的解决交通问题，这时问题的归属有了一个新的定义——交通流量预测。目前应用效果最佳的是短时交通流量预测，基于预测结果形成的智能交通系统也可以更好的疏导交通，提升交通的畅通性。

现代智能交通系统的核心是短时交通流量。交通流量预测的长时性预测和短时性预测。长时性的交通预测在目前迅速发展的交通状况中并不能很好的解决人们所面临的交通拥堵问题，所以研究的重点被人们放在了短时性预测。

短时性交通预测的实现随着计算机硬件的发展和人们对深度学习的研究的深入而得以实现。

## 国内外研究现状[王晓杰]

交通里预测方法多采用以随机过程、统计学以及时间序列方法、滤波方法和人工智能方法。各国学者和研究者尝试将各种方法或算法用于交通流预测，按预测使用的方法理论用于道路预测的模型或方法有很多，根据预测路段的数量交通流预测模型可以分为单路段预测和多路段预测，也叫单点预测和多点预测。预测单路段或者单个检测器的交通流量的模型或方法有历史均值法、回归分析法、时间序列法、Kalman滤波法、神经网络模型。历史均值法、时间序列法、卡尔曼滤波法和神经网络既可以进行单点预测也可以实现多点预测，用于多点交通流预测的方法还有动态交通分配法和非参数回归法。

比较早用于交通流预测的方法是由**R.E.Kalman**提出的一种线性滤波方法一卡尔曼滤波法。卡尔曼滤波法模型中的参数和权重调整机制按照线性无偏均方误差来计算输出与输入的偏差，模型不断的迭代调整整个系统的权重使得误差达到最小或者理想状态。**Okutani**使用卡尔曼滤波系统进行交通流预测时，在模型中加入了预测路段和相邻路段之间的影响关系，提高了模型的准确率;杭明升将Kalman滤波模型用于预测车辆在高速公路上行程时间;通过研究，杨兆升发现一个路段的流量会受路段的上游流量的影响，当前路段的流量会影响其下游路段的交通路况，在使用卡尔曼滤波算法构建交通流预测模型考虑了上下游路段对预测路段的影响；朱中把交通流量和路段的旅行时间结合起来构建数学空间模型，将卡尔曼滤波算法用于模型求解上，比较使用不同的观测矩阵元素对模型预测的影响，发现使用流量比值效果较好。邵春福等也使用卡尔曼滤波求解根据交通流量构建的空间模型来将流量转化为旅行时间。

**Yang** 用卡尔曼滤波算法求解噪声协方差矩阵和状态转移矩阵组成的广义交通流预测模型。

卡尔曼滤波模型是一种线性估计模型，比较适合数据集是线性的和变化比较规律性，并且时间间隔最好是大于5分钟。经网络模型能够识别线性数据，能从数据中自动找寻出交通流的特征。神经网络学习是黑箱式学习，可以不用完全了解隐含层的学习机制，可直接用于数据的分类预测。

**Dougherty**等基于BP神经网络无需输入大量参数的优势，用来预测短期交通流量，机动车车速，车量占有率，不过预测效果没有达到理想要求。**Van Lint**将神经网络加入递归模式对高速公路旅行时间进行预测，预测准确率较高。**Park B**基于相同的数据集分别构建反向传播神经网络模型和径向基神经网络模型实现交通流预测，分析两个模型的实验结果，发现向基神经网络模型复杂度较小并且准确率较高。**Rilett L R**等将整个路径化分为多个路段，利用谱分析神经网络构建网络模型，然后分别对单个路段的行程时间进行预测，将预测结果整合成路径的行程时间。尹宏宾提出的模糊神经网络算法在预测交通流上具有无参数，误差小，算法计算时间短的优点。

**Ahmed和Cook** 发现交通流数据具有时序的特点，于是基于高速公路和车辆占有率数据构建时间学列模型对其进行分析研究，并在模型中引入了博克思一詹金斯法方法，要比单纯使用时间序列模型对流量和占有率预测分析具有较高的精确度。**Nihan N. L.和Holmesland K.O**用同样的方法预测了一个路段的交通流量。**Lingras Pawan**等通过研究交通流数据，发现某一时刻的交通流量与上一时刻或者几个时间间隔以前的交通流量有一定的关系，因此，将数据集按照道路类型进行分类，分别分析这些不同类型的数据的时序特征，利用时间序列模型可以分析这些数据实现交通流预测。国内学者**朱顺应、徐今强，谭国贤，翁小雄**等也将交通流数据在时间上的关联性的特征应用与交通流预测上。**韩超**对构建好的时间序列模型进行参数估计，使用的是带遗忘因子的最小二乘法，误差函数使用的是线性最小方差函数。宗春光等利用了相空间重构理论对交通流数据分析研究，并且从局部分析数据集特征，结合计算出的相空间重构参数进行建模实现交通流的预测。**黄**使用非线性混沌预测模型对交通流进行预测，首先提取时间相关性的交通流数据之间的相关信息，然后用混沌理论对已经构建好的预测模型中的预测方程参数进行求解，并通过误差分析对整个模型的参数进行调整优化。唐明等利用交通流数据的模糊特征和分形特征将混沌理论和分形理论应用于构建交通流预测模型中，研究分析交通流数据的流量变化和数据集与时间有关的特征。田晶基于同一数据集构建BP神经网络模型和混沌时间序列模型，分析两个模型在预测同一交通流数据集时的优缺点，发现混沌时间序列方法具有更高的实用性。蒋海峰提出对局部分析交通流数据并部分加权的方法构建模型实现交通流预测，能更好的适应各种类型的数据集。

**Davis和Nihanl**基于在拥挤和非拥挤之间的交通流数据对非参数回归方法和时间序列算法进行了比较，发现前一种方法的预测效果更好。**Stephenark**在非参数回归方法中加入多个变量来对处于多种交通状态下的交通流进行预测。**Turochy R E**研究了非参数回归在交通领域的应用，使用不同类型的非参数回归方法进行实验比较其区别。**达庆东**利用非参数回归模型建立了基于交通流量和车辆密度的预测模型，基于车流量和车辆速度的预测模型，基于车辆速度与车辆密度的预测模型。**宫晓燕**没有单独使用非参数回归算法进行交通流预测，而在基本的模型基础上结合基于密度的K搜索算法，并且数据预处理上使用的方法是基于动态的聚类算法和散列函数，改进后的模型可以同时满足高准确率和低复杂性。

目前，各国的研究者和专家依旧对短时交通流预测进行着深入研究，因为短时交通流的预测不仅可以解决交通问题，也是完善智能交通系统的前提条件。除了分析单一的路段或者路段上的某个截面的交通流外，研究者还对某个区域的路网的交通状况进行研究。对于路网的研究一般使用的是数据挖掘的方法，数据挖掘的方法主要是基于路网中相互连通的路段之间的相关性进行整体分析研究。根据路网内各个道路相互连通的性质，分析有检测器的路段的数据去推断无检测器路段的数据的特征。不过数据挖掘的方法只考虑了路段与路段之间的参数数量关系，忽略了路段间的拓扑结构。

**杨兆升**将城市道路分为有检测器和无检测器两类，有检测器的路段能够采集到交通流数据，根据路段与路段之间的相关性，用逐步回归分析法对与路段之间的相似性进行分析，寻找与无检测器道路相似的有数据的路段，利用其数据进行分析研究去预测无数据路段的交通状况。张赫等也是利用交叉路口的检测器数据，分析研究发现交叉路口间的车道之间是具有相关性的，使用聚类分析和逐步回归法预测无数据的交叉路口上的车道的拥堵情况。**宫晓燕**等用关联规则算法挖掘路段流量之间的相关性，利用某个路段的流量去推测其他与己知路段具有相关性的路段的流量信息。不过这个方法精确度不高。

面对互联网的蓬勃发展，信息快速的传播使得信息量成指数形式增加，我们可以轻易获得全球世界各国的各种信息，大部分数据人们都可以共享。随着智能交通系统对一个城市的影响越来越大，交通领域各个路段每时每刻的交通信息不断被采集保存在交通数据库中。传统的交通流预测方法都是人工从数据库中寻找数据集的特征，并且适用于小数据样本的预测。面对海量的交通流信息，研究者受人工神经网络的启发，提出来深度学习的概念。深度学习是能够处理海量数据，无需对输入数据进行复杂的预处理的一种包含多个隐含层和感知器的一种学习结构。深度学习模型的网络结构都是非线性的，对非线性知识具有比较好的学习能力，对将要学习的知识通过学习后将其特征进行抽象表达，能够自动学习和记忆学习过的数据中的特征。

**Hinton**等首次使用深度学习信念网络模型识别数据图形；**Kuremoto**将波尔兹曼机模型与深信度网络模型组合。**Huang**等利用深信度网络对流量数据进行多任务输出的交通流预测。将自编码算法应用到交通预测领域。

## 本文主要工作

XXXXX

## 文章组织结构

XXXX