

我阅读了交通流量预测相关的一些综述等资料，初步整理了如下内容：

1. 数据集
2. 研究现状

整体来说，通过初步资料阅读，我初步了解了交通流量预测领域的研究现状，认识了主流的几种技术方向(CNN, RNN, GNN 等)，并认识了常用相关数据集。

## 数据集

---

参考 [TKDE2020综述](#) 的分类标准，主要数据集可以分为两类：

- 固定式采集数据集(point data)：安装在固定地方的探测器所采集的数据
- 移动式交通数据(trajjectory data)：GPS 等收集的车辆轨迹信息

此外，可能还需要一些辅助数据信息：

- 交通网络数据：探测器的分布图(欧氏空间网格或无向加权图等)
- 气候数据、日期(节假日)数据、事件(如车祸)数据等

列出部分常用的数据集：

- [PeMS](#) (Caltrans Performance Measurement System)  
研究最广泛的数据集，由加利福尼亚州主要公路的上万探测器收集，每半分钟采集一次，包含容量、速度、交通流量等多种数据  
有多个子集广泛用于论文中，包括 PeMS-BAY、PeMSD3、PeMSD4、PeMSD7、PeMSD8 等，可以参考 [这篇综述](#)
- [METR-LA](#) (Metro Traffic Los Angeles)  
洛杉矶公路网，207 个探测器，5 分钟间隔收集数据
- [Seattle Loop](#)  
西雅图 4 条路数据，323 个探测器，5 分钟间隔收集数据，2015 年 1 月数据
- [SZ-Taxi](#)  
深圳罗湖区 156 条路的数据，15 分钟粒度，2015 年 1 月数据
- [Beijing Traffic](#)  
北京市 3126 个路段在 2022 年 5-7 月的 5 分钟粒度的数据
- [Q-Traffic](#)  
北京 2017 年 4-5 月一万多个路段每 15 分钟采样一次的百度地图数据

其他数据集：参见 [paperwithcode 网站：Traffic Prediction Task](#)、[这篇综述](#)、[TKDE2020综述](#)、[这篇综述](#)。

## 研究现状

---

交通流量问题的定义，参见 [TKDE2020综述](#) 的描述(其他综述类似)：

- 以历史的交通流量数据为输入，学习未来交通流量的预测问题。
- 流量主要受时间、空间因素影响，还受到天气节假日等因素影响。

最初使用经典统计模型如 ARIMA 和传统机器学习方法如 kNN，但是现在的主流已经是深度学习。由于主要影响因素是时空，因此：

- 针对时间因素，主要使用 RNN，如 LSTM 处理
- 针对空间因素：
  - 建模为欧氏平面网格图，使用 CNN 处理
  - 建模为非欧空间，使用 GNN 处理

通常将二者结合为混合神经网络，来解决该问题

对 CNN, RNN 技术和详细论文列表可以参考 [TKDE2020综述](#)，GNN 参考[这篇综述](#)

CNN 有包括但不限于下面的思路：

- 交通流读入可以建模为图像，每个像素点对应一个交通密集的地区，因此可以用图像识别技术，即把区域网格化，像素值是如车辆数目，不同时间即像素值不一样
- 考虑到数据集空间可能连城一条线，可以使用一维 CNN 或二维 CNN(一维时间、一维空间)
- 经纬作二维，然后引入时间为第三维度做三维 CNN
- .....

CNN 的部分实现论文如下：[经典二维空间](#)、[CNN+GRU](#)、[三维 CNN 结合经纬+时间信息](#)、[局部 CNN+LSTM](#)、[一维CNN+2层LSTM](#)

RNN 有包括但不限于下面的思路：

- 将 RNN 的输出与 CNN 的输出使用 FNN 融合在一起
- 使用流水线，先处理时间再处理空间或反过来
- 修改模型内部结构，如融入图卷积等，使其结合空间特征
- .....

RNN 的部分实现论文如下：[LSTM](#)、[LSTM+GRU](#)、[二维LSTM变体](#)、[考虑缺失数据的 LSTM-M](#)、[Meta-RNN](#)、[注意力+LSTM](#)、[SAE+LSTM](#)

除了 CNN, RNN 外，还有包括但不限于这些新的方向：

- 使用 GNN 相关技术代替或结合 CNN，占据主要数据集的 SOTA 和前列
- 使用注意力机制代替 RNN 来获取时间信息
- 使用 Transformer 等 encoder-decoder NLP 技术

列举部分 [paperwithcode](#) 上的 SOTA 论文如下：

- METR-LA 数据集：
    - [STGM](#) 注意力捕获时空依赖+GNN(空间处理)+改进CNN(时间处理) 2023 (SOTA)
    - [Traffic Transformer](#) (Transformer) 2020
    - [SLCNN](#) (CNN+图) 2020
    - [STAEformer](#) (Transformer) 2023 (也是 PeMSD7 的 SOTA)
    - [D2STGNN](#)、[STEP](#)、[MegaCRN](#) (GNN) 2022
    - [DCGCN](#) (GNN) 2023; [DCRNN](#) (有向图扩散+CRNN) 2018
  - PeMSD8 数据集：
    - [STWave](#) (图+注意力) 2023 (SOTA)
    - [STD-MAE](#) (两个 autoencoder 分别处理时空) 2023
- PeMS-BAY、PeMS07、PeMSD04、PeMS04 四个数据集的 SOTA

- [PDFormer](#) (Transformer) 2023
- [DDGCRN](#) GNN 2023
- 其他数据集:
  - SOTA PeMS08 [HTVGNN](#) (注意力(时间)+改进GNN(空间)) 2024
  - SOTA NE-BJ [RGDAN](#) (图扩散+注意力) 2024
  - .....