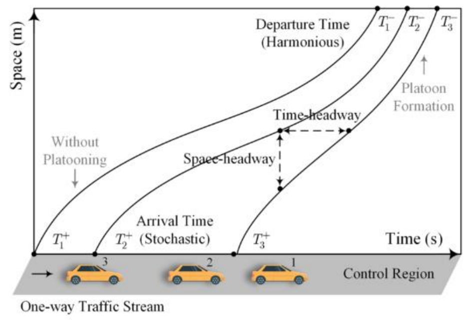
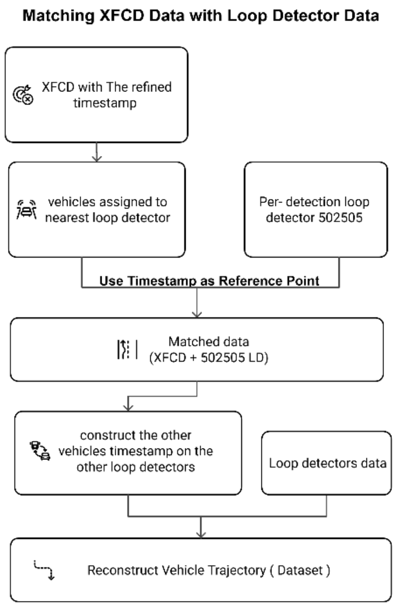
Identification of Vehicle Trajectories using Machine Learning Approaches

Master‘s Thesis of Mohamed Ibrahim

Mentoring:
M.Eng. Athanasia Karalakou
Dr.-Ing. Arslan Ali Syed

Univ.-Prof. Dr.-Ing. Klaus Bogenberger
Chair of Traffic Engineering and Control





Vehicle trajectories are essential for traffic state estimation, planning, and autonomous systems. However, full-resolution trajectories are hard to obtain due to the high cost of data acquisition. This work addresses this by combining Extended Floating Car Data (XFCD) with loop detector timestamps using machine learning.

Problem Statement

<u>XFCD Limitations:</u> While Extended Floating Car Data provides rich vehicle dynamics (e.g., speed, acceleration, lane position), its timestamps are often coarse or irregular due to rounding (hourly intervals). This temporal misalignment complicates integration with infrastructure-based loop detectors, which offer precise, fixed-interval data but lack spatial granularity.
<u>Loop Detector Limitations:</u> LD provides aggregated traffic metrics (flow, speed) at specific highway points but cannot capture continuous vehicle trajectories.

Objective

Develop a machine learning framework to fuse XFCD and LD datasets, refine timestamps, and reconstruct high-resolution vehicle trajectories for applications in traffic management, congestion analysis, and road safety.

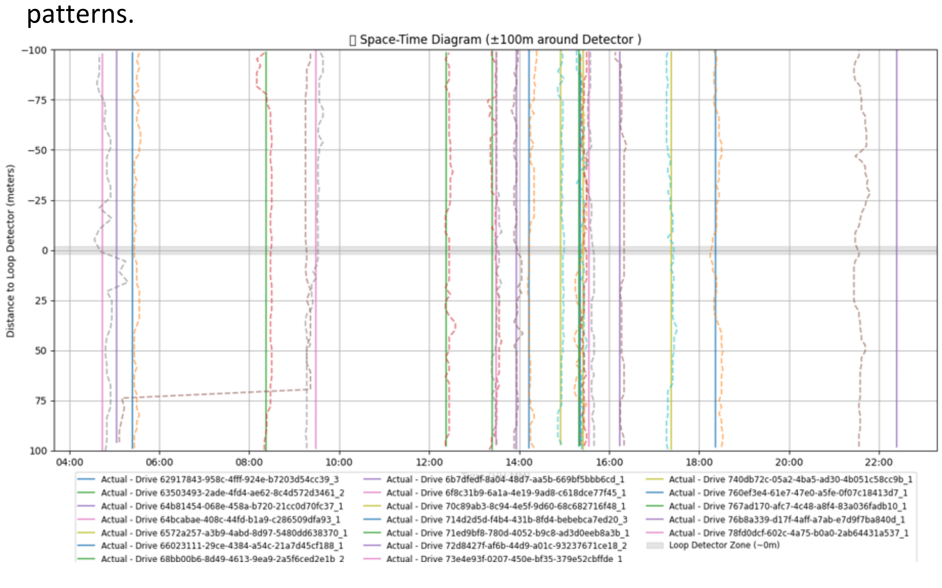
Preprocessing Pipeline

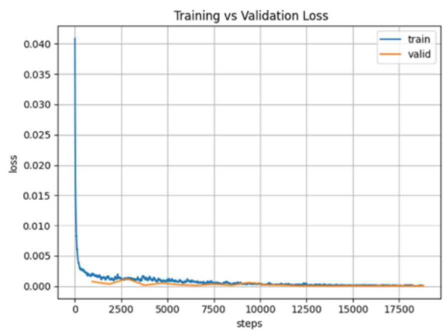
1. Directional Filtering: Split XFCD into northbound/southbound trajectories using latitude trends (monotonically increasing/decreasing in latitude) and discarded unreliable heading angle data due to GPS noise and Lane changes.
2. Outlier Removal: Eliminated GPS drifts and Ramps.
3. Timestamp Reconstruction: Combined TAbsMs (absolute trip start time) and TRelMs (relative time since trip start) to derive refined timestamps.
4. Spatial Matching: Aligned XFCD with LD data within ±100 meters of detector locations to focus on the vehicle measurements.

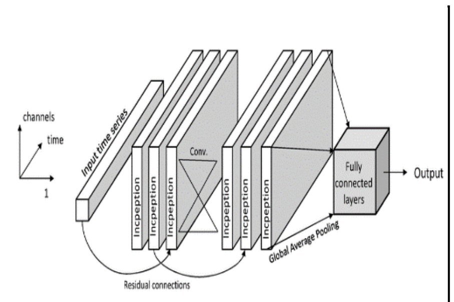
Neural Network Models

<u>InceptionTime CNN:</u> Parallel 1D convolutions with kernel sizes to capture multi-scale temporal patterns.

<u>LSTM:</u> Sequential recurrent network for temporal dependency learning.







Metric	InceptionTime CNN	LSTM
MAE (seconds)	524.59 (~8.7 minutes)	1,200 (~20 min)
RMSE (seconds)	911.86 (~15.2 minutes)	1,800 (~30 min)
R² Score	0.90	0.60
±60-min Accuracy	99%	64%

The study successfully addressed the temporal misalignment between sparse loop detector (LD) timestamps and dense Extended Floating Car Data (XFCD) by developing a machine learning framework that enables continuous vehicle trajectory reconstruction. By leveraging the InceptionTime CNN architecture, the model achieved superior performance over traditional LSTMs, reducing timestamp prediction errors by 56% (MAE: 524.59 seconds) and explaining 90% of the variance in the data (R^2 = 0.90). The framework’s scalability allows deployment across large-scale traffic networks, making it a practical tool for Intelligent Transportation Systems (ITS). These advancements directly support congestion analysis, real-time incident detection, and adaptive traffic control by providing high-resolution trajectory data.
To further advance the capabilities of this framework, future efforts will focus on exploring hybrid LSTM-CNN architectures, combining the temporal dependency learning of LSTMs with the multi-scale pattern recognition of CNNs to enhance timestamp refinement