Neural Network-Based Trajectory Prediction with Criticality Metrics: Feature Analysis and Model Optimization

Master‘s Thesis of Mykyta Yevtushenko

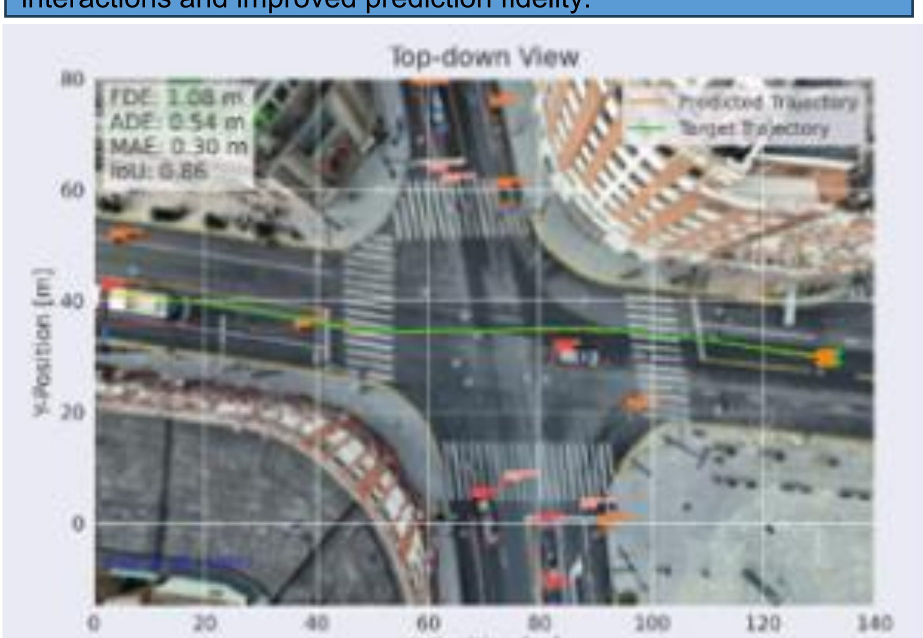
Mentoring:
M.Sc. Johannes Lindner, M.Sc. Evald Nexhipi

External Mentoring:
M.Sc. Azur Causevic (BMW AG)

Driving Simulator at BMW AG

Approach

A **deep-learning trajectory prediction** approach based on an **attention-based encoder** and **LSTM decoder** processes **historical motion data** to generate multi-step **future trajectory predictions**, yet current models lack consideration of **interactions with other traffic participants**, reducing accuracy. Using data from a large **urban intersection in Sichuan**, originally containing only **x/y coordinates**, **frame IDs**, and **track IDs** for **960 participants**, including **400 motorbikes**, this work expands the feature space with **criticality metrics (CMs)** commonly applied in **autonomous driving**. These metrics describe how **dangerous** a trajectory is by accounting for surrounding agents and were selected based on **scenario coverage**, **sensitivity**, and **real-time suitability**. The integrated CMs include **Headway**, **Time-Headway**, **Time-To-Collision**, **Time-To-Zebra**, **Delta-V**, **Collision Probability**, and **IUTQ**. Additional interaction-focused features—such as **occupied spaces**, **Top-K participants** (distance- and TTC-based), and **relative motion descriptors** from the **ego-vehicle perspective**—enable richer modeling of potential interactions and improved prediction fidelity.



Example of a good prediction. Green = Ground Truth
Orange = Prediction

Introduction

Simulations play a critical role in the **vehicle development cycle**, yet a persistent **simulation–reality gap** limits their effectiveness. Realistic modeling of **road agents**—particularly **motorbikes** and **scooters**—is essential, as **road-user interactions** in large urban intersections are highly complex. Existing **VRU-specific research** is limited, especially for **motorbikes**, with key challenges including the **lack of datasets** and the need for **domain-specific models**. Prior work, such as the deep-learning approach introduced by **Causevic et al.**, uses relatively simple datasets containing many VRUs, highlighting the need for more refined methods. This work analyzes an existing **dataset** to identify and integrate new **behavioral features** into a **neural-network-based prediction framework**, aiming to improve **trajectory prediction accuracy** across varying **prediction horizons**. Addressing the challenge of accurately forecasting scooter trajectories enhances the **realism of traffic simulations**, enabling more reliable evaluation of **driver interactions** and **advanced assistance systems**.

Experiment	Comment
01	Basic NN + Basic features by [2]
02	Basic NN + Full feature set (basic and newly computed)
03	Basic NN + Full feature set + manually reduced CMs
04	Basic NN + PCA-reduced full feature set
05	Basic NN + Full feature set + PCA reduced CMs
06	Basic NN + feature selection via BO - basic configuration from step 0
07	Basic NN + feature selection via BO - basic configuration from step 150
08	Basic NN + feature selection via BO - 20 random initial samples
09	NAS and feature selection at the same time
10	just NAS and basic features
11	feature selection, after performing NAS in 10

Experiment overview

10	**3.79854**	**6.78283**	**0.45403**

Performance for best experiment (ADE | FDE | IoU)

Results

Experimental results indicate that introducing additional **features** did not enhance **prediction accuracy**, as the added inputs introduced **noise**, leaving the **basic feature set** as the most robust. A **neural architecture search (NAS)** using this simpler input configuration achieved the best performance, though notable **inaccurate predictions** still occurred. Adjusting **input–output sequence lengths** showed only minor influence, confirming that **data quality** and **model architecture** matter more than sequence tuning. Several factors limit performance: **static criticality metrics**, **tracking errors**, **missing or imbalanced data**, lack of **lane-level context**, simplified **motorcycle dynamics**, and insufficient **behavioral diversity**, all hindering generalization. Overall, the model lacks the **spatio-temporal reasoning** required for reliable real-world deployment, and current **prediction errors** remain too large for practical use.