

ADL - CLEANING THE INPUT DATA FOR TRAINING SELF DRIVING CARS

HELENA SOKK
HENRILAATS
KRISTJAN LAID

INTRODUCTION

Tartu University's Autonomous Driving Lab has managed to get the driving recordings obtained with the self-driving Lexus on Rally Estonia tracks. The problem they are facing is that currently the dataset contains information that they don't need or that would influence the outcome. Therefore the dataset is unusable at the moment. That's the part where our team comes in to help them clean the dataset into an usable one.

DATA

The data was stored in a virtual machine as rosbag files, from which we had to extract it from. It contained 100GB of data. We used topics on position, steering, speed, time, turn signal state.



CLEANING DATA

Removing the parts where the car was not moving. We dropped the rows in data frame where the vehicle speed was zero.

Removing the parts where a turn signal was used. We dropped the rows in data frame where the turn signal was zero.

Removing the unused parts. Dropped topics like yaw, roll, pitch, left/right side images.

Converting the data

Steering angle. Converted from radians to degrees.

Coordinates. Created a function to convert into L-EST97 coordinate system.

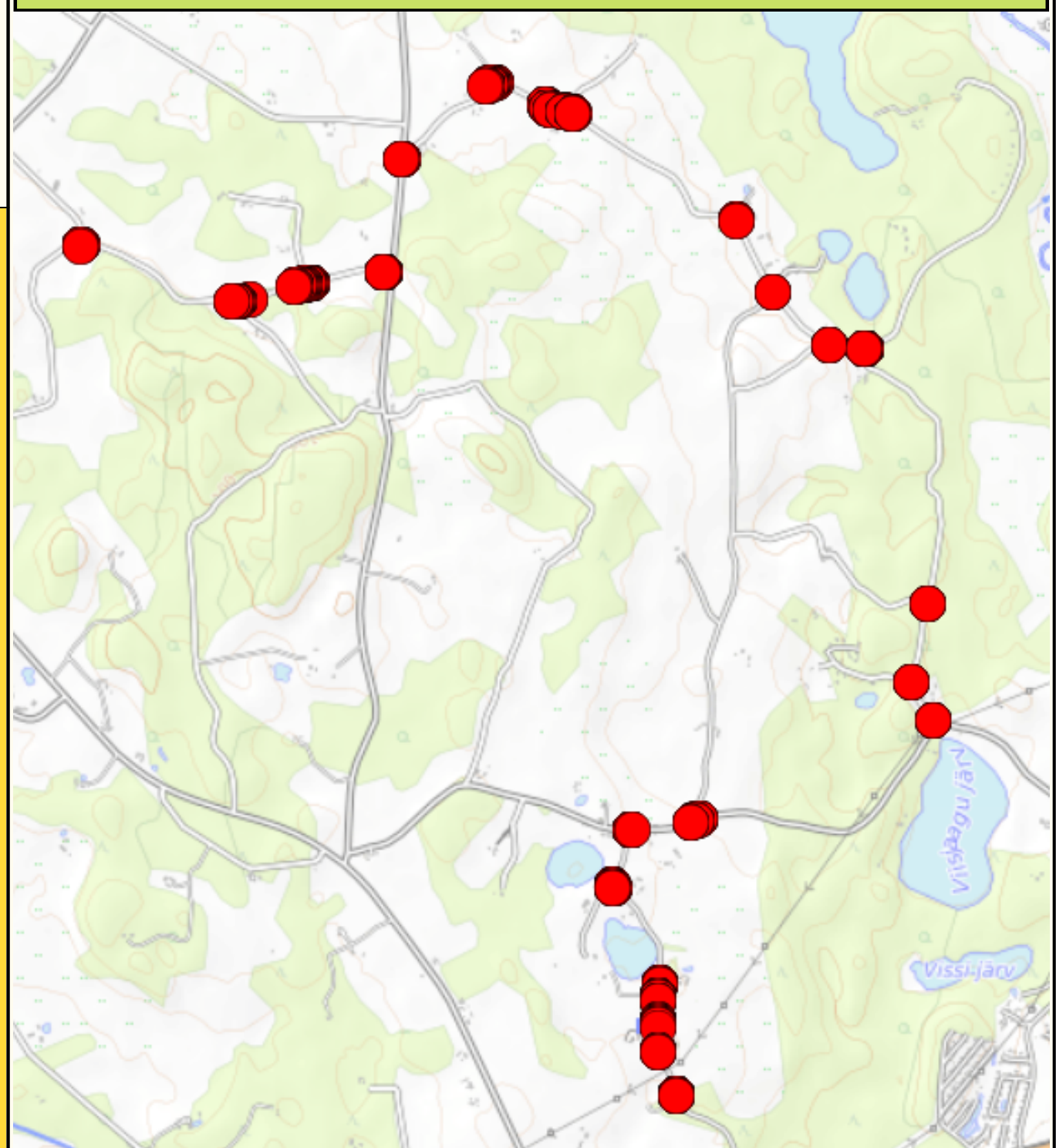
Time. Converted days/hours/minutes into seconds.

Speed. Converted from m/s to km/h.

Separating data

At first the data was separated into turns and straights using steering angle and coordinates. Then it was separated into "just follow the road" and complicated intersections datasets using coordinates.

In the picture we have automatically mapped out the intersections.



[Link to Github repository.](#)