Motion Prediction for Autonomous Vehicles

DA Project

Temwirk
Vignesh K Kumar
Tarun Gupta J
Arun Srinivasan P
Chandratop Chakraborty

Introduction

Our task is to build motion prediction models for self-driving vehicles. Model which can predict the movement of traffic agents around the Autonomous Vehicles such as cars, cyclists, and pedestrians etc. We are required to predict how these different agents move in an Autonomous Vehicle’s environment.

What is an Autonomous Vehicle?

An autonomous vehicle, or a driverless vehicle, is one that is able to operate itself and perform necessary functions without any human intervention, through ability to sense its surroundings.

An autonomous vehicle utilises a fully automated driving system in order to allow the vehicle to respond to external conditions that a human driver would manage.

Dataset

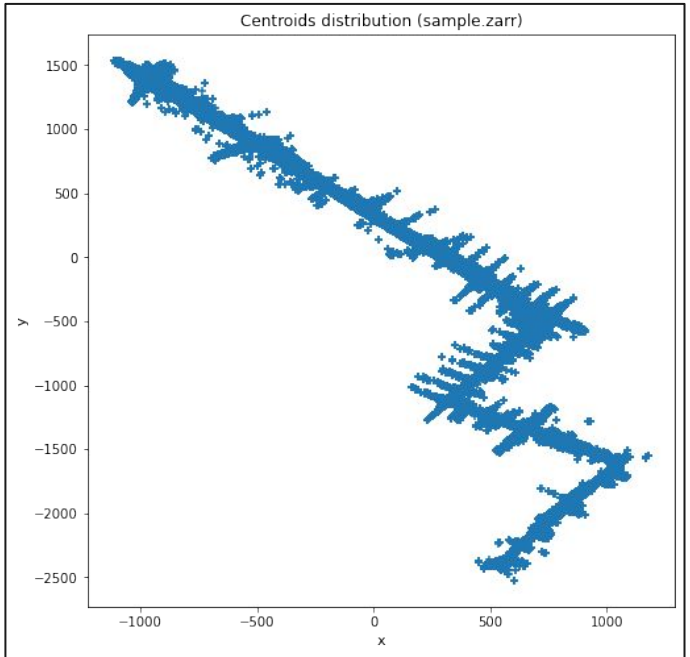
This dataset includes the logs of movement of cars, cyclists, pedestrians, and other traffic agents encountered by Lyft's autonomous fleet. These logs come from processing raw lidar, camera, and radar data through our team’s perception systems and are ideal for training motion prediction models. The dataset includes:

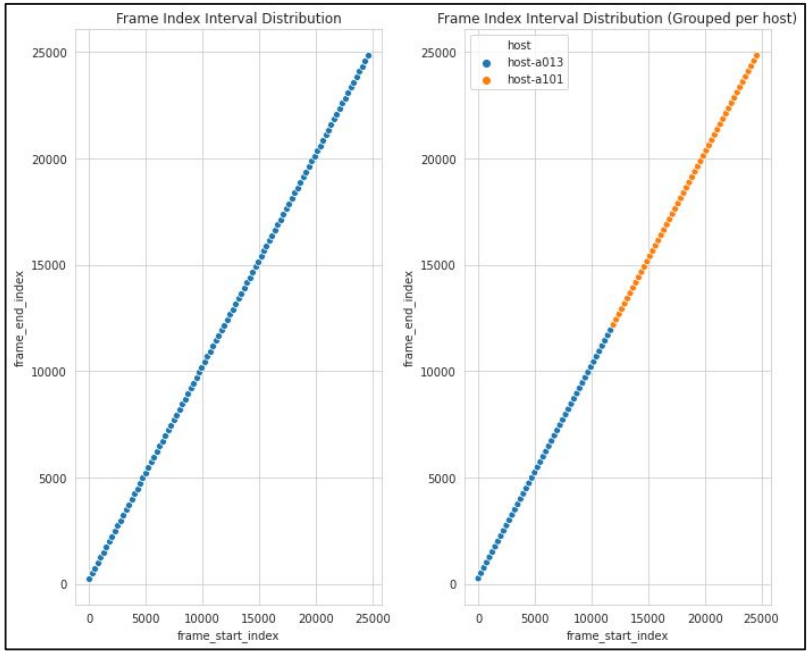
- 1000+ hours of traffic agent movement
- 16k miles of data from 23 vehicles
- 15k semantic map annotations

Selective Visualizations

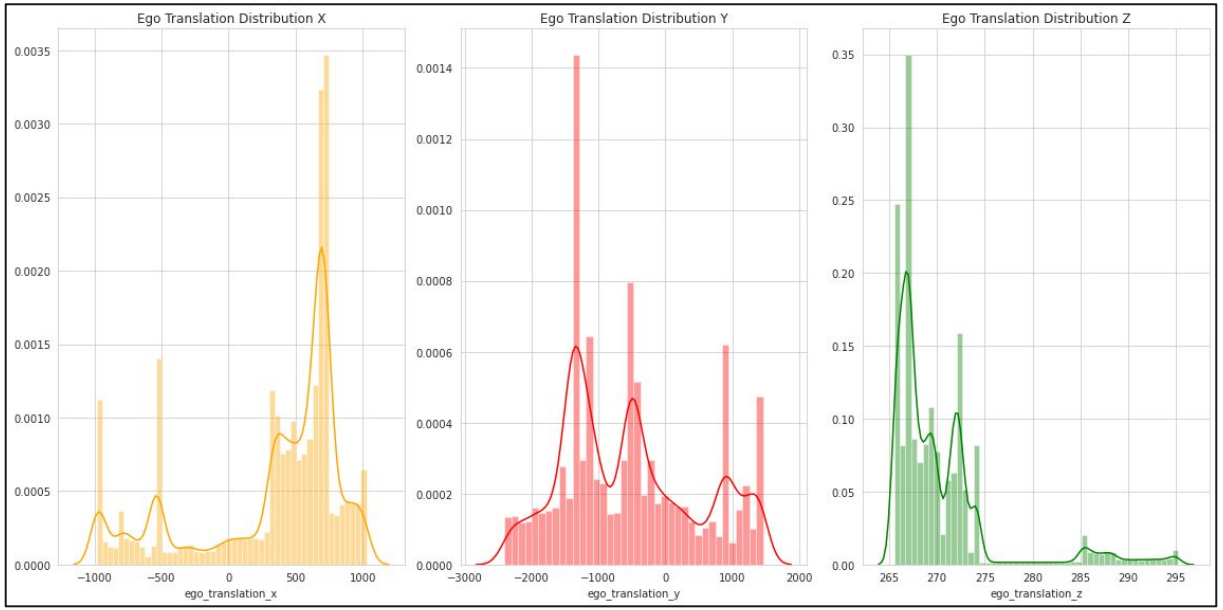
- Our entire code base is hosted on the kaggle cloud notebooks.
- The dataset can be accessed through the cloud link and the code can be executed as well.
- The entire code takes around 2 hrs to finish execution.

Visualizations - 1

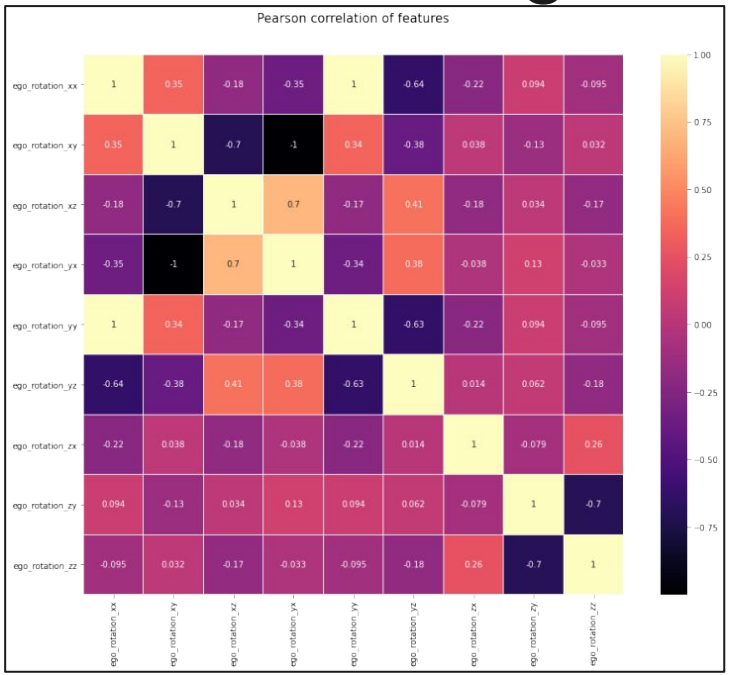


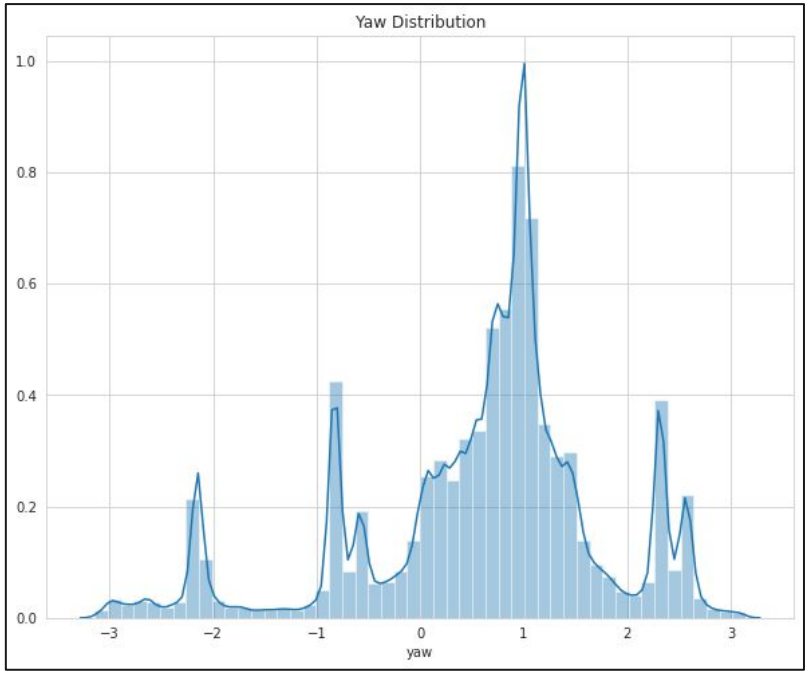


Visualizations -2

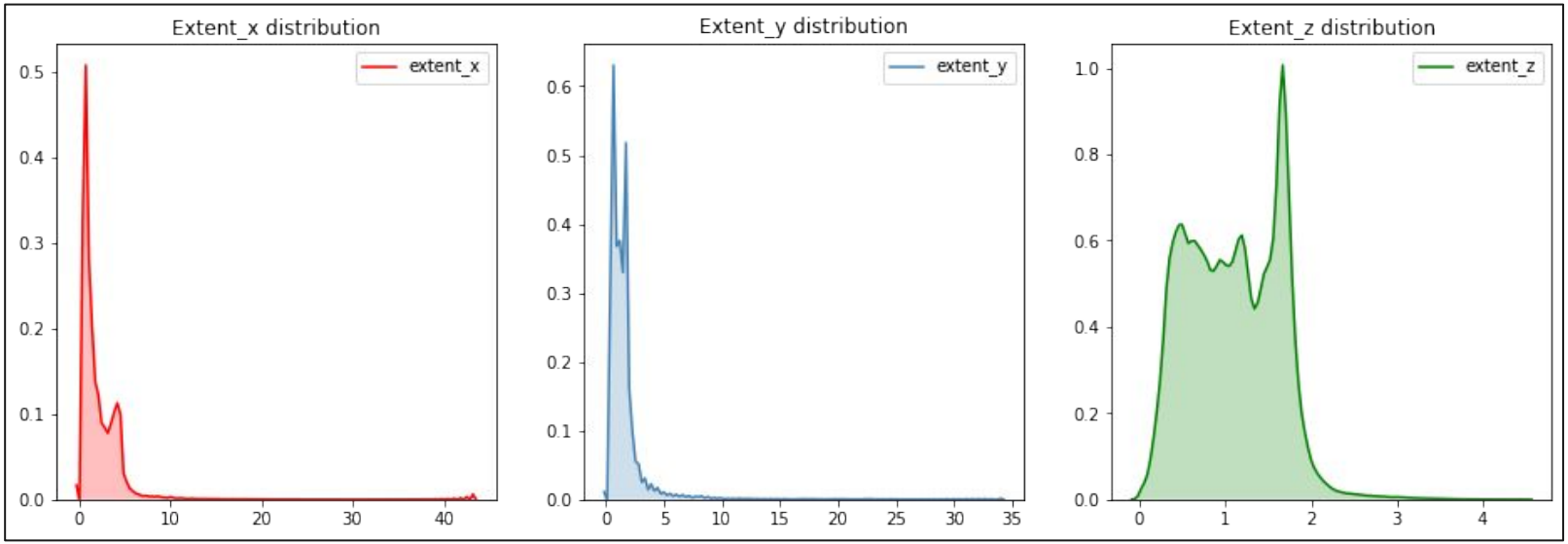


Visualizations -3



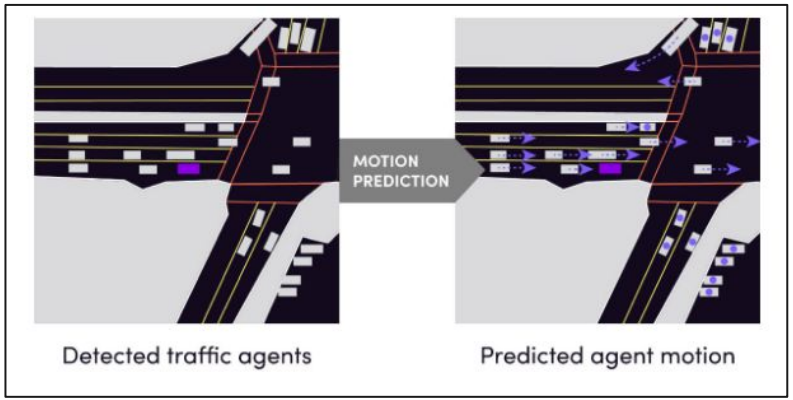


Visualizations -4



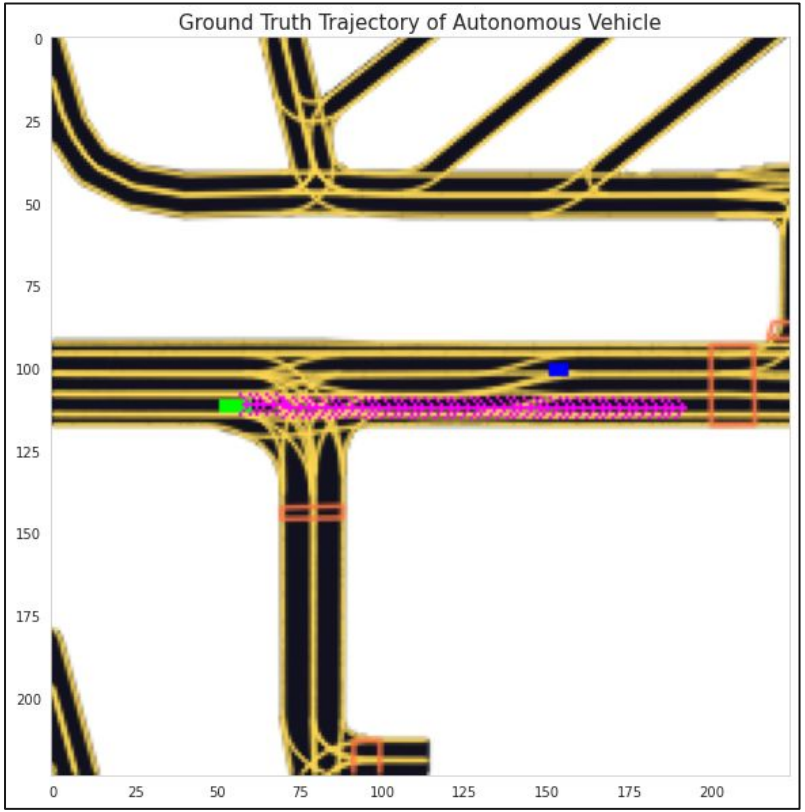
What are we predicting?

Our task in the competition is to predict the motion of external objects such as cars, cyclist, pedestrians etc in order to assist the self-driving car. We have to predict the location of objects agents in the next 50 frames.



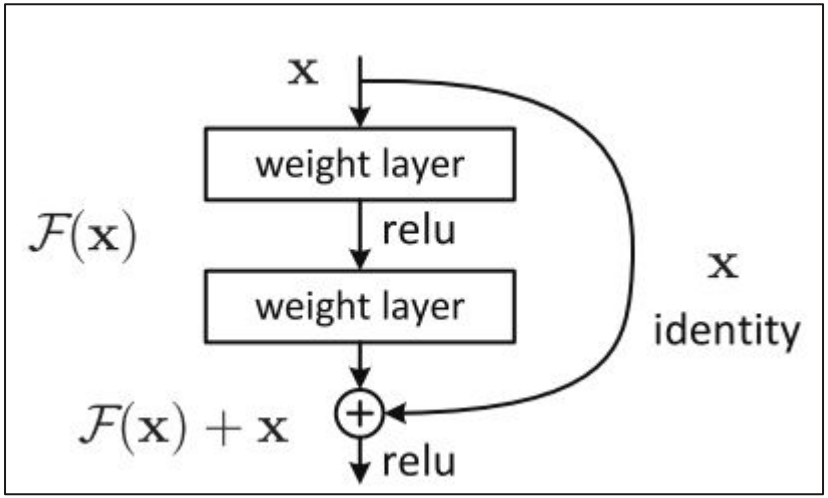
Ground Truth

The pink path represents the ground truth that has to be predicted.



Approach

We have accomplished the given task utilizing ResNet architecture. It seemed to give us the best results among the other approaches we tried.



Evaluation of the output

We calculate the negative log-likelihood of the ground truth data given the multi-modal predictions.
Loss: Our task is to minimize this in order to get a better score

$$
\begin{aligned}
L &= -\log p(x_{1,\dots,T}, y_{1,\dots,T} | c^{1,\dots,K}, \bar{x}^{1,\dots,K}_{1,\dots,T}, \bar{y}^{1,\dots,K}_{1,\dots,T}) \\
&= -\log \sum_k e^{\log(c^k) + \sum_t \log \mathcal{N}(x_t | \bar{x}^k_t, \sigma=1) \mathcal{N}(y_t | \bar{y}^k_t, \sigma=1)} \\
&= -\log \sum_k e^{\log(c^k) - \frac{1}{2}\sum_t (\bar{x}^k_t - x_t)^2 + (\bar{y}^k_t - y_t)^2}
\end{aligned}
$$

Our Submission

We got a score of 23.610 based on the evaluation metric in the previous slide. This score was in the 300s rank range. The score was obtained through the kaggle platform as well.

Software Used:

1) L5Kit
2) Python
3) Numpy
4) Pandas
5) Seaborn
6) Pretty Table
7) Matplotlib
8) PyTorch

THANK YOU