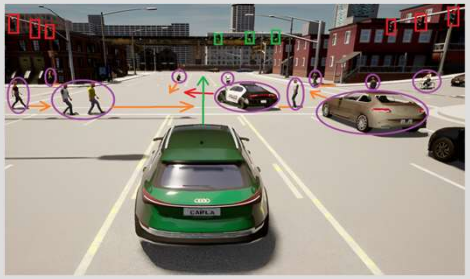
Graph-based Prediction and Planning Network (GP3Net) for Scalable Self-Driving in Dynamic Environments Using Deep Reinforcement Learning

Jayabrata Chowdhury[1]*, Venkataramanan Shivaraman[2]*, Suresh Sundaram[1], P.B. Sujit[2]
[1]Indian Institute of Science, Bangalore and [2]Indian Institute of Science Education and Research, Bhopal

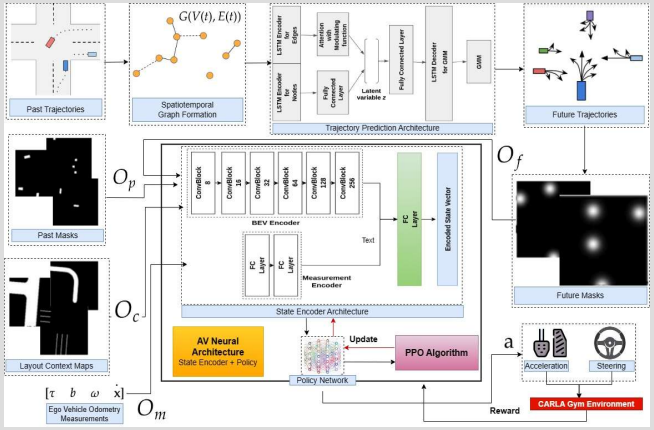
Motivation



Challenge:
- Navigate safely in a **non-stationary** environment
- Expert design rules and imitation learning on driving data are not scalable and suffers distribution shift
- Recent methods should integrate traffic intentions with safe exploration for motion planning better

Solution:
- Model how the driving context can change in near future with uncertainties
- Enhance reinforcement learning based driving capability with better exploration of the driving scenario

Our Method



Step 1:
- Dynamic graph: each traffic participant is a node, connected by their interactions
- LSTM networks to analyse individual movements (position, velocity, acceleration)
- Edge encoder LSTM network analyses connections between nodes
- CVAE to generate multimodal predictions for future trajectories with GMMs
- Account for the **prediction uncertainty using 2D Gaussian patches** centered on likely positions.

$$g = exp\left(\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}\right)$$

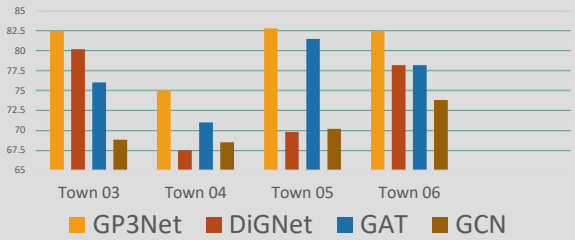
Step 2:
- AV learns by doing, adjusting its decision based on **dense rewards**
- The feedback loop drives quick and efficient learning

$$R = r_{route} + r_{halt} + r_{vel} + r_{pos} + r_{hd} + r_{act} + r_{term}$$

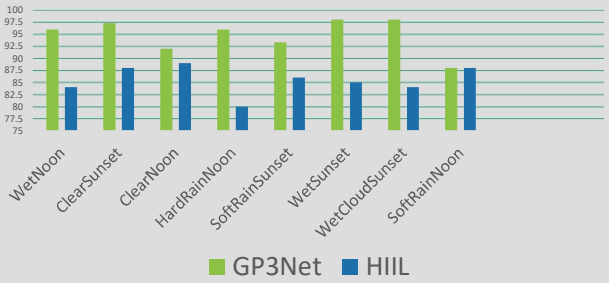
- Use a modified PPO algorithm with an entropy term for safe exploration

Performance

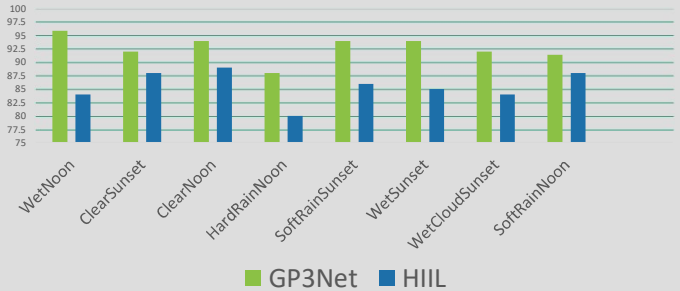
Success Rate (SR)



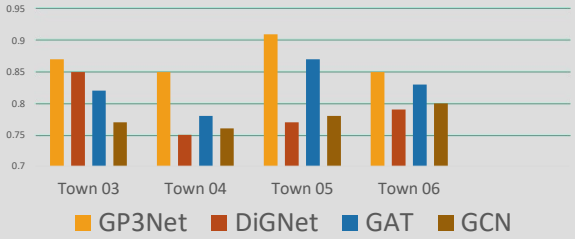
Success Rate (SR) Town 01

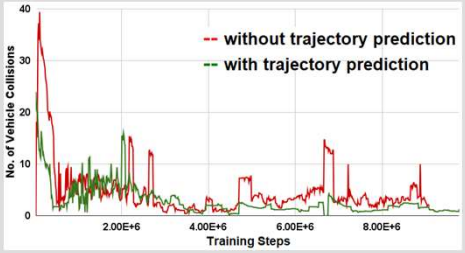


Success Rate (SR) Town 02



Driving Score (DS)





- Outperforms imitation learning models on standard CARLA benchmarks (urban, highway, mixed)
- Adapts to different weather conditions, completing routes with fewer infractions
- The average advantage: 3.85% in SR and 8% in DS with a low standard deviation of 1.5%

- Focus on initial phase: mastering route-following and avoiding collisions in various situations
- GP3Net paves the way for robust and adaptable AV navigation

- Predicting future trajectories enhances safety measures in dynamic environments
- The videos of GP3Net can be found by scanning the QR code.