Trajectory Prediction for Real-world Autonomous Driving

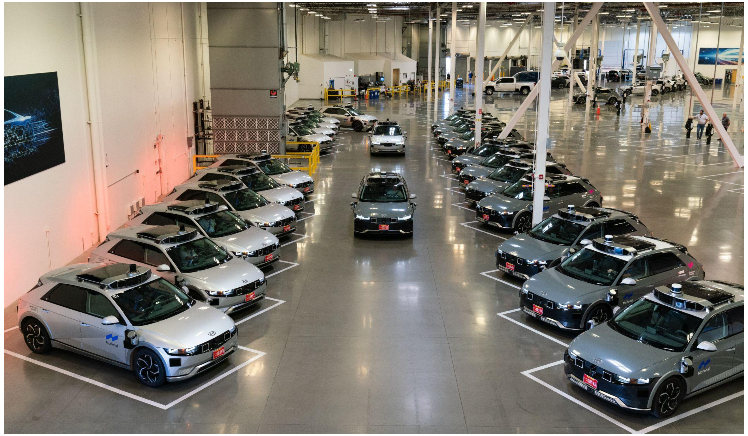
Henggang Cui
Prediction Team Lead
Motional

Motional's AD Technology

Motional has been at the forefront of autonomous driving innovation for over a decade:

- nuTonomy (2013) -> Aptiv (2017) -> Motional (2020)
- Launched the world's first robotaxi pilot in Singapore in 2016.
- Launched a public robotaxi service on the Lyft network in Las Vegas in 2018.
- First fully driverless vehicles operation on public roads in 2020.
- Now operating autonomous ride-sharing and delivery services on the Uber and Lyft networks in multiple cities.

Motional's all-electric IONIQ 5 robotaxi



Motional's robotaxi fleet at the company's Technical Center in Las Vegas

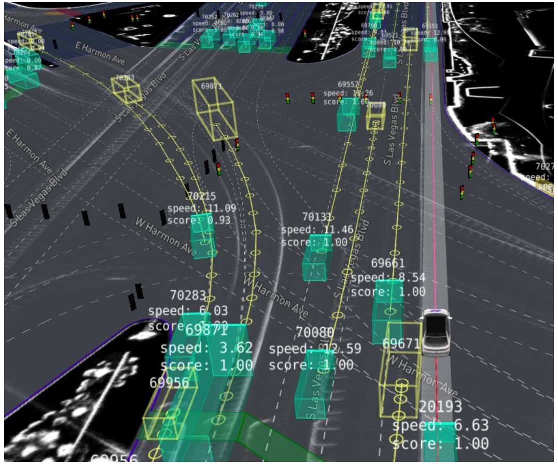
Trajectory Prediction Research and Real-World Autonomous Driving

Trajectory prediction is a critical part of Motional’s autonomous driving technology

- Trajectory prediction (a.k.a. motion forecasting) predicts the future behaviors of the other road users
- Trajectory prediction is also a popular research area
  - Public benchmarks (e.g., Argoverse and WOD) has tremendously boosted trajectory prediction research
- However, there are a few important gaps between trajectory prediction research and real-world autonomous driving
  - Map compliance
  - Compute latency

Rank	Participant team	brier-minFDE (K=6) (↑)	minFDE (K=6) (↑)	MR (K=6) (↑)	minADE (K=6) (↑)	DAC (K=6) (↑)	minFDE (K=1) (↑)	MR (K=1) (↑)	minADE (K=1) (↑)	La su at
1	SEPT-iDLab (SEPT)	1.6820	1.0566	0.1032	0.7282	0.9922	3.1777	0.5154	1.4412	23
2	QCNet-AV1 (QCNet)	1.6934	1.0666	0.1056	0.7340	0.9887	3.3420	0.5257	1.5234	6 r
3	ProphNet (ProphNet)	1.6942	1.1337	0.1101	0.7623	0.9893	3.2628	0.5261	1.4910	9 r
4	FFINet (FFINet)	1.7286	1.1213	0.1124	0.7606	0.9875	3.3611	0.5431	1.5327	10 ag
5	USTB_yhf (distprob_62)	1.7305	1.1293	0.1084	0.7996	0.9903	3.3579	0.5390	1.5388	28
6	Polkach (VI LaneIter)	1.7313	1.1057	0.1065	0.7709	0.9897	3.2849	0.5275	1.5181	1 y
7	Wayformer	1.7408	1.1616	0.1186	0.7676	0.9893	3.6559	0.5716	1.6360	1 y

Argoverse Motion Forecasting leaderboard

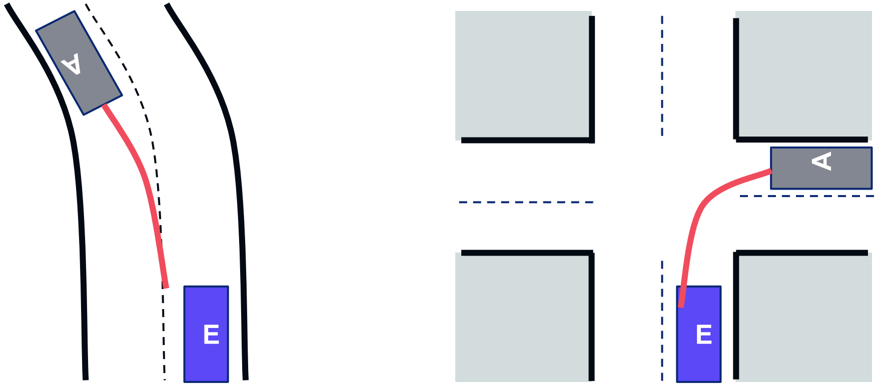


Real-world trajectory prediction task

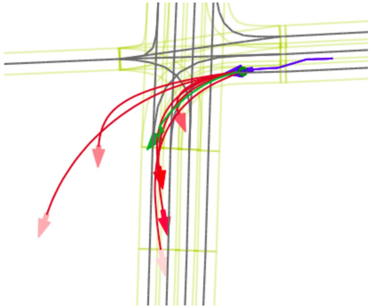
Map Compliance

Map compliance is a characteristic that the predicted trajectory needs to be compliant with the map

- Non-map compliant trajectories (going off road or into other lanes) lead to bad and even dangerous ego behaviors
- Unfortunately, map compliance is often neglected by the standard motion forecasting benchmarks
  - Not by displacement error or miss rate if the displacement to ground-truth is small
  - Not by the drivable area compliance (DAC) metric if the trajectory is still in drivable area
- The benchmarks' top-6 evaluation metrics encourage diverse prediction but don't penalize the bad predictions



Non-map-compliant trajectory predictions



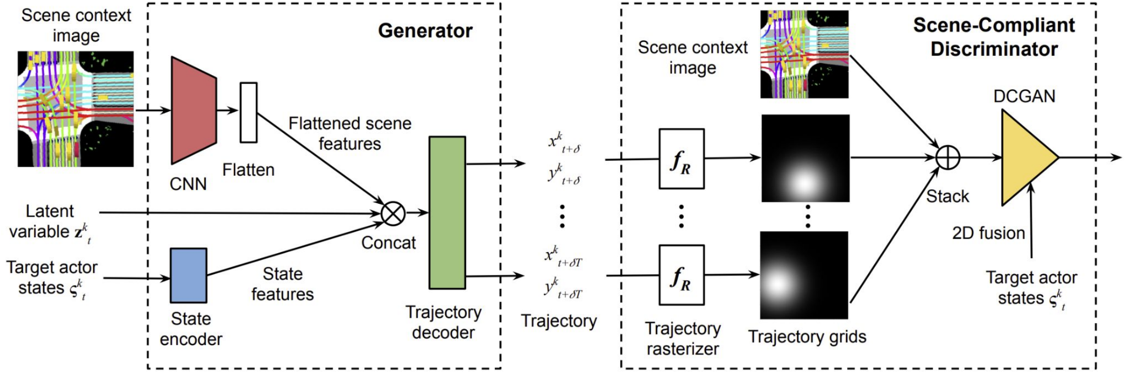
Non-map-compliant predictions are not penalized by the standard benchmarks

Prior Works on Map Compliance

There are some prior works on improving prediction map compliance with rasterized map

- With rasterized map [1], it is a hard computer vision problem to predict trajectories that follow the rasterized lanes
- Scene-compliant GAN in [2], ellipse loss in [3], and spatial grids in [4] helps map compliance with rasterized map

Rasterized map [1]

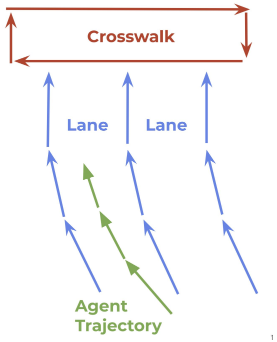


Scene-compliant GAN in [2]

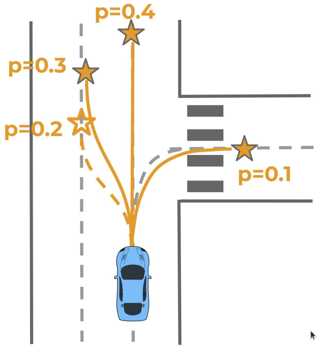
[1] Cui et al. Multimodal Trajectory Predictions for Autonomous Driving using Deep Convolutional Networks. ICRA 2019
[2] Cui et al. Improving Movement Predictions of Traffic Actors in Bird's-Eye View Models using GANs and Differentiable Trajectory Rasterization. KDD 2020
[3] Cui et al. Ellipse Loss for Scene-Compliant Motion Prediction. ICRA 2021
[4] Ridel et al. Scene Compliant Trajectory Forecast with Agent-Centric Spatio-Temporal Grids. RAL 2020

Prior Works: Goal-Based Prediction

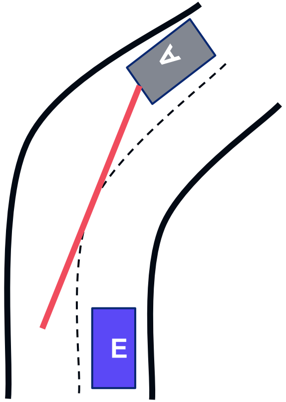
- Goal-based prediction (e.g., TNT [2]) with vectorized map (e.g., VectorNet [2])
  - Sample candidate goals (trajectory endpoints) on lane centerlines
  - Classify goal probabilities using goal features
  - Predict trajectories conditioned on goals
- Improves map compliance because the goals can be sampled from lane centers
- The trajectory endpoint will be map compliant, but the rest of the trajectory can still be non-map-compliant



Vectorized map with VectorNet [2]



Goal-based prediction in TNT [1]

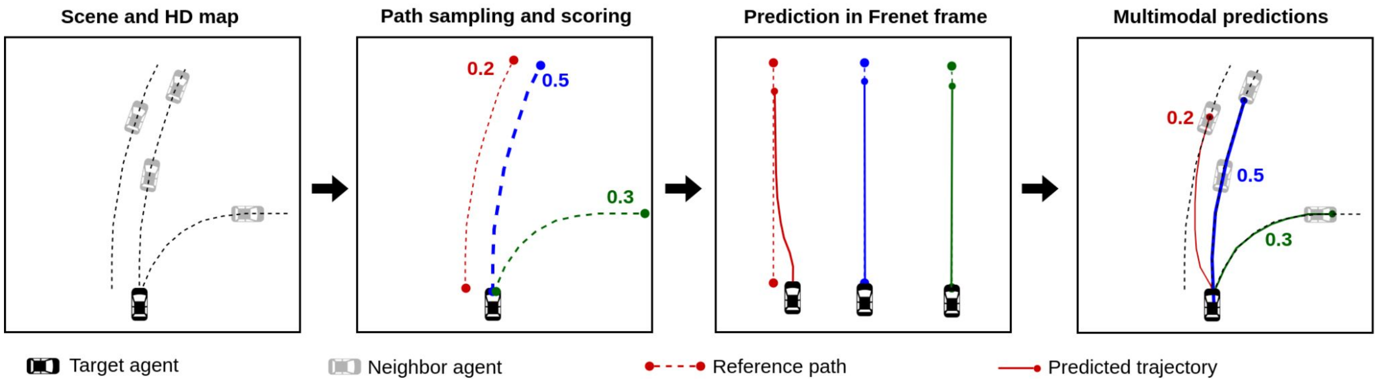


Non-map-compliant prediction even when the goal endpoint is on lane center

[1] Zhao et al. TNT: Target-driveN Trajectory Prediction. PMLR 2021
[2] Gao et al. VectorNet: Encoding HD Maps and Agent Dynamics from Vectorized Representation. CVPR 2020

Our Solution: Path-Based Prediction [1]

We propose path-based prediction [1] that significantly improves prediction map compliance

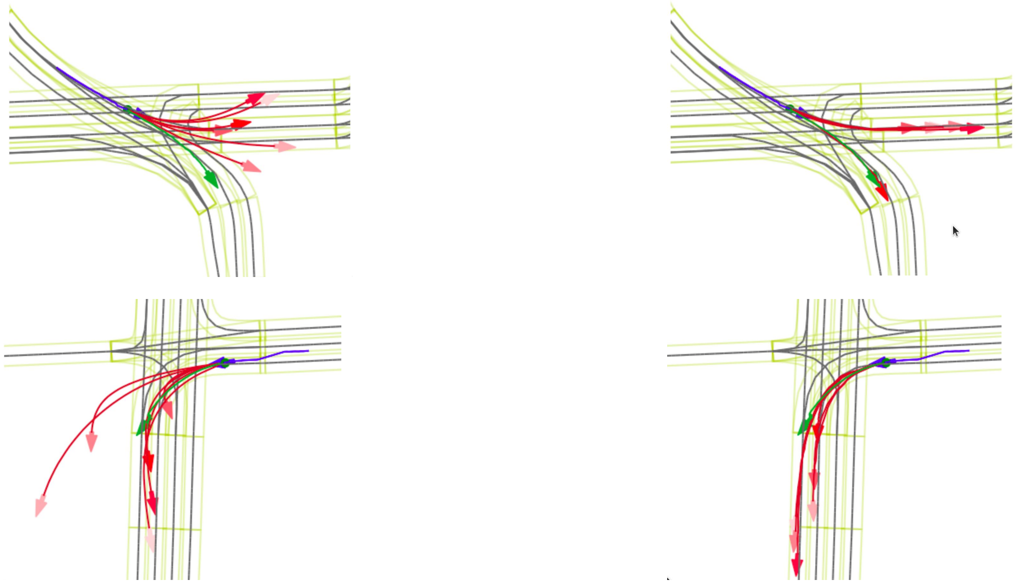


- Path-based prediction
  - Sample candidate **reference paths** based on lane centerlines
  - Classify **path probabilities** using **path features**
  - Predict trajectories conditioned on **reference paths** in **Frenet frame**
- Advantages over goal-based prediction
  - Path features instead of goal features
  - Prediction in the Frenet frame

[1] Afshar et al. PBP: Path-based Trajectory Prediction for Autonomous Driving. arXiv 2309.03750

Our Solution: Path-Based Prediction [1]

Path-based prediction significantly improves map compliance on Argoverse



Blue: history
Green: ground truth
Red: top-6 predictions

Left: traditional multimodal prediction   **Right: path-based prediction**

[1] Afshar et al. PBP: Path-based Trajectory Prediction for Autonomous Driving. arXiv 2309.03750

Our Solution: Path-Based Prediction [1]

Path-based prediction has the best map compliance on Argoverse leaderboard

- DAC (drivable area compliance) is the only metric on Argoverse leaderboard that measures the map compliance
  - Higher number is better
- Path-based prediction (PBP) has the best drivable area compliance among all state-of-the-art methods

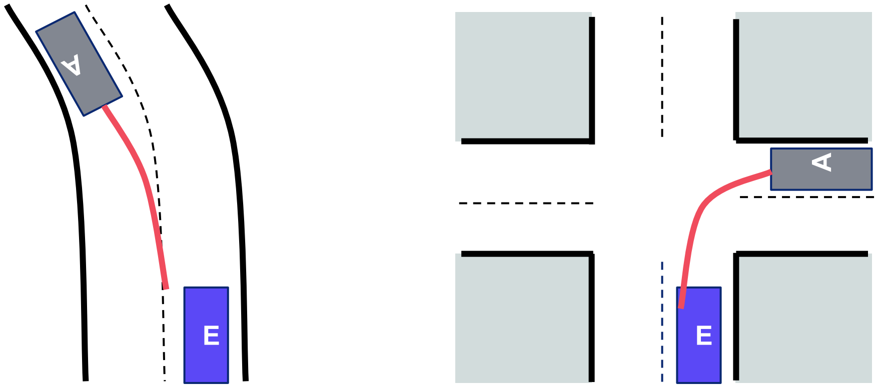
TABLE II: Comparison to the state-of-the-art models on the Argoverse leaderboard

Model	minADE_1	minFDE_1	MR_1	minADE_6	minFDE_6	MR_6	DAC
TNT [43]	2.174	4.959	0.710	0.910	1.446	0.166	0.9889
DenseTNT [16]	1.679	3.632	0.584	0.882	1.282	0.126	0.9875
GoHOME [14]	1.689	3.647	0.572	0.943	1.450	0.105	0.9811
PRIME [32]	1.911	3.822	0.587	1.219	1.558	0.115	0.9898
HiVT-128 [46]	1.598	3.532	0.547	0.773	1.169	0.127	0.9888
MultiPath++ [33]	1.623	3.614	0.564	0.790	1.214	0.132	0.9876
DCMS [40]	**1.477**	**3.251**	**0.532**	0.766	1.135	0.109	0.9902
Wayformer [23]	1.636	3.656	0.572	0.767	1.162	0.119	0.9893
QCNet [45]	1.523	3.342	0.526	**0.734**	**1.067**	**0.106**	0.9887
PBP (Ours)	1.626	3.562	0.535	0.855	1.325	0.145	**0.9930**

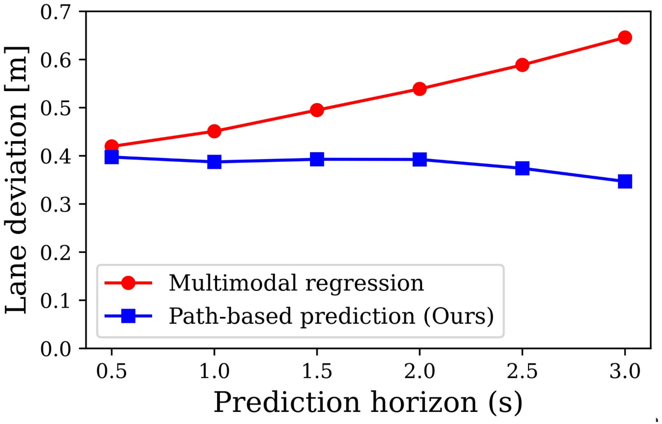
[1] Afshar et al. PBP: Path-based Trajectory Prediction for Autonomous Driving. arXiv 2309.03750

Measure Map Compliance with Lane Deviation Error

- DAC (drivable area compliance) only penalizes predictions going outside the drivable area
- We propose the **lane deviation error** metric
    - Measures as the distance from each prediction waypoint to the closest lane *that follow the same direction*
    - Measured at each prediction horizon
    - Averaged among all published trajectories
- **Path-based prediction achieves significantly lower lane deviation error compared to the multimodal baseline, especially at longer horizons**



Non-map-compliant predictions in drivable area

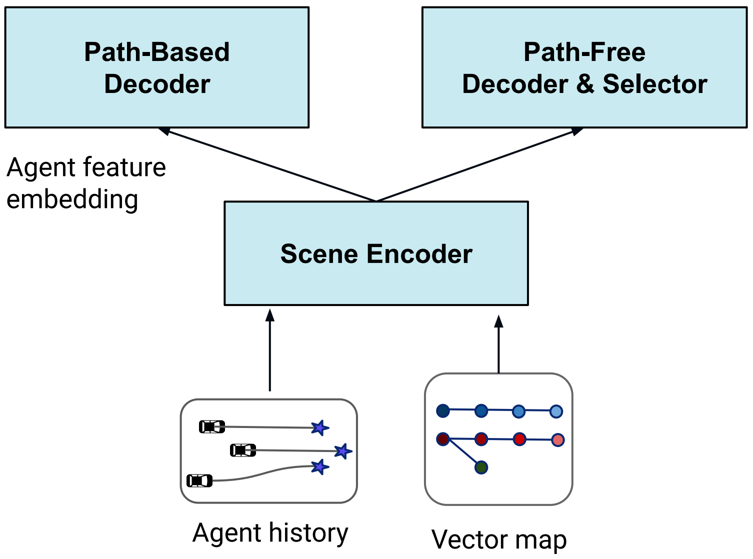


Path-based prediction vs. multimodal baseline on lane deviation error

How about the Non-Map-Compliant Agents?

Handle non-map-compliant agents with a path-free decoder

- Sometimes agents go offroad, and there can be areas with no map annotations (e.g., private driveways)
- To handle non-map-compliant agents, we pair path-based prediction with another path-free prediction decoder
- And we train a classifier to select between the two decoders for each agent
- The two decoder heads share the same encoder backbone so it introduces very little runtime overhead



Compute Latency

Compute latency is important for real-time systems

- In a 10 Hz autonomous vehicle stack, the trajectory prediction module needs to run under 100 ms at 99 percentile
- The standard trajectory prediction benchmarks don't measure compute latency
    - Participants run their solution on the test set locally and submit their prediction results to the leaderboard
- Many methods on the leaderboard are computationally expensive
    - Per-agent prediction that runs scene encoding for each agent
    - Large ensemble models with high number of replicas

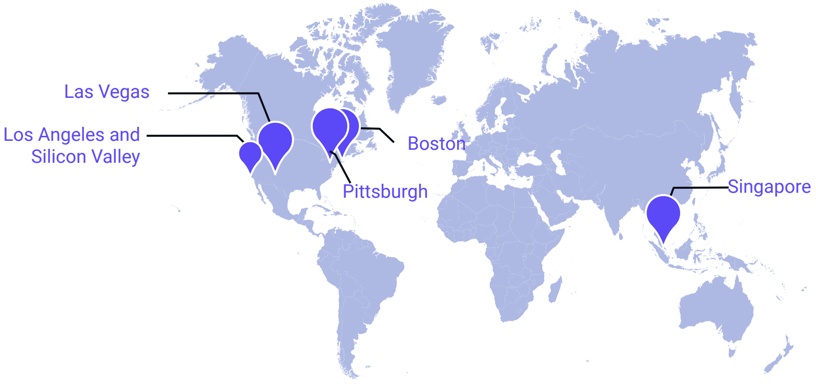
- **We adopted an efficient ego-centric prediction architecture that predicts for all agents in the scene in one pass**
    - **99 percentile inference latency is only ~40 ms on a whole scene**

Conclusions

We discuss two important gaps between trajectory prediction research and real-world autonomous driving

- Map compliance
  - We propose a path-based prediction approach that significantly improves map compliance over SOTA
  - We propose a new lane deviation error metric to measure map compliance
- Compute latency
  - We use an efficient ego-centric architecture to predict all agents in the scene in one pass with under 40 ms

We are hiring! https://motional.com/careers



Many office locations and support fully remote!