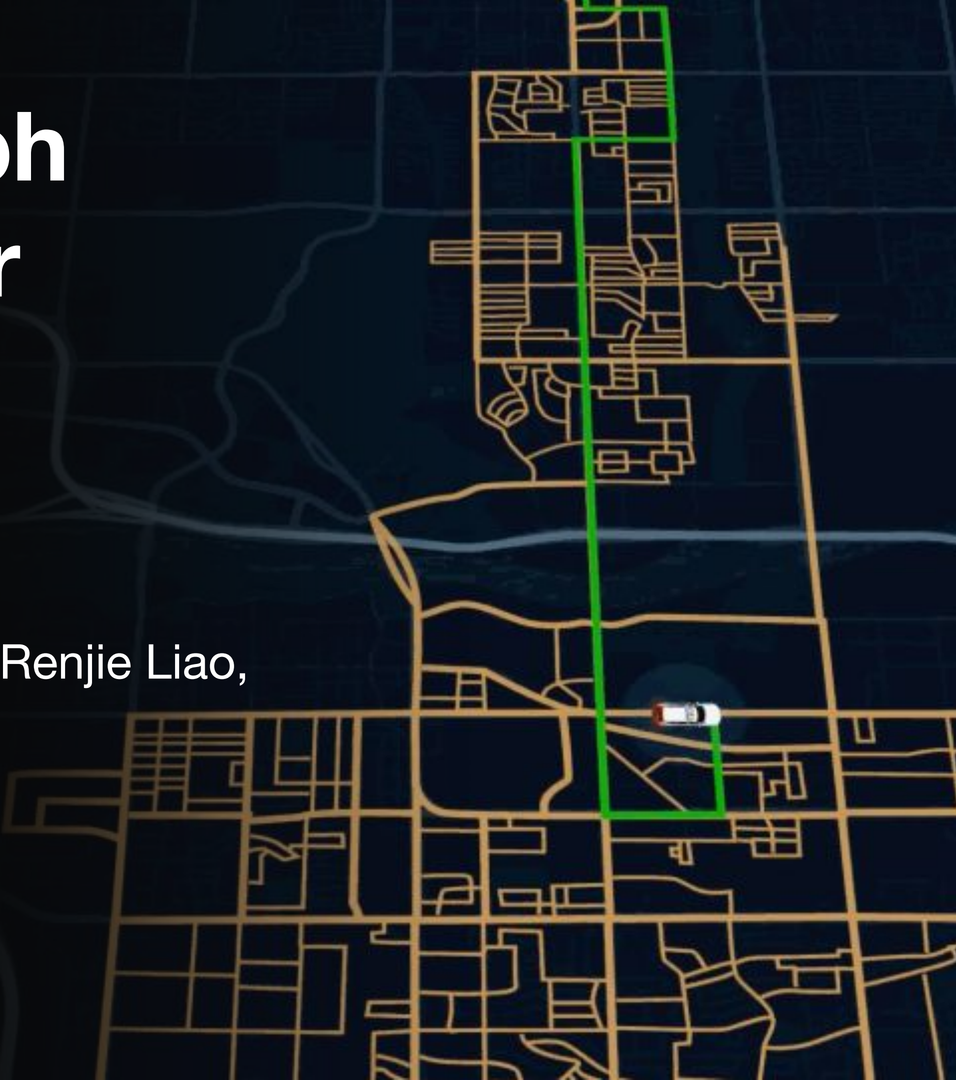
Learning Lane Graph Representations for Motion Forecasting

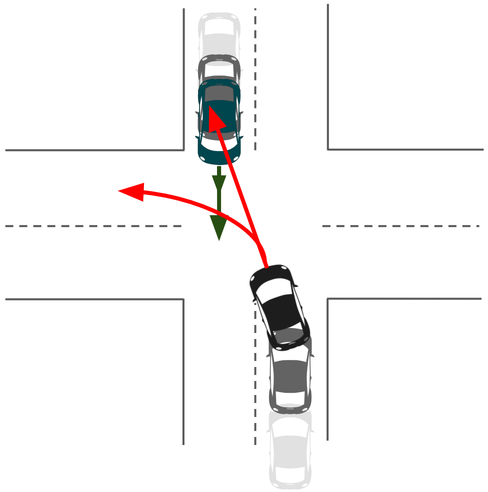
Ming Liang, Bin Yang, Rui Hu, Yun Chen, Renjie Liao,
Song Feng, Raquel Urtasun

UberATG
UNIVERSITY OF TORONTO



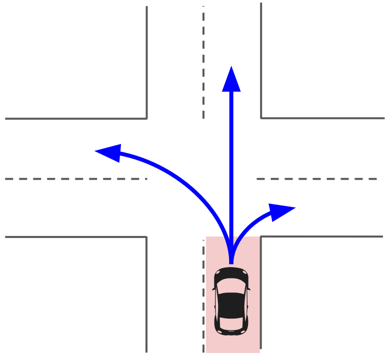
HD Maps for Motion Forecasting

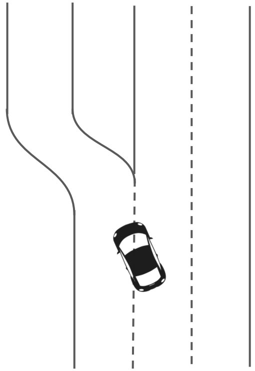
- Motion forecasting predicts future trajectories of actors given their past states
- HD maps provide useful clues for motion forecasting
  - Behaviors of traffic agents mostly depend on the map topology
  - Interactions of agents are conditioned on maps



Related Work: Heuristics

- Rule-based vehicle & lane association
- Multi-model trajectories with follow-lane assumption
- Drawbacks:
  - The vehicle & lane association is error-prone
  - Cannot generalize to complex driving behaviors (e.g., lane change)

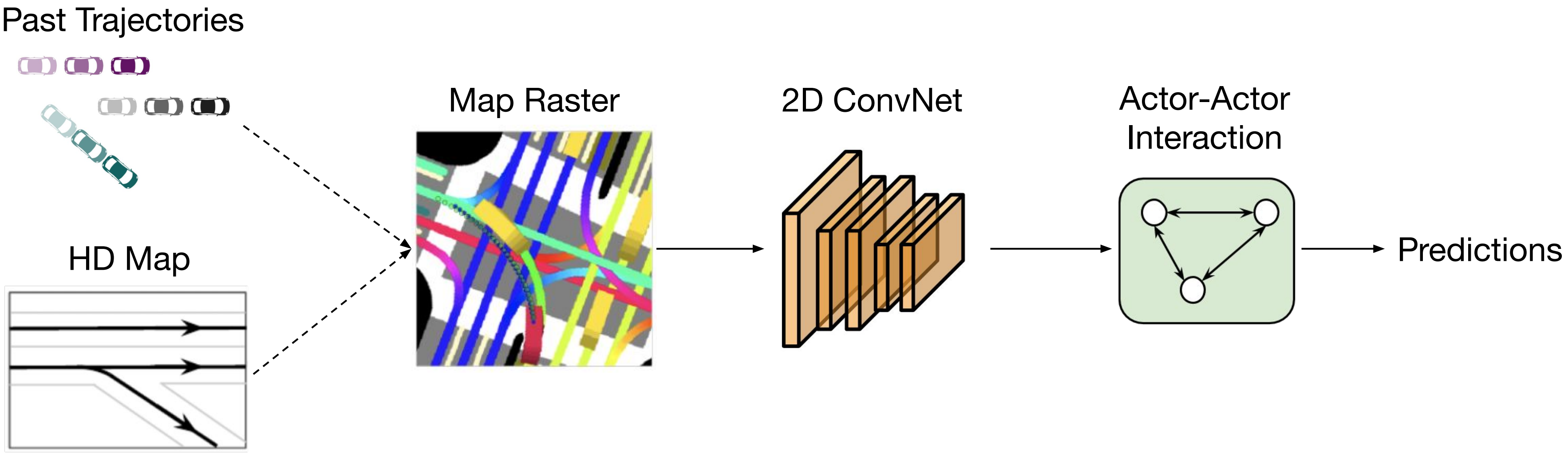




[1] Making Bertha Drive—An Autonomous Journey on a Historic Route. [J. Ziegler, et al. 2014]

Related Work: Raster Images

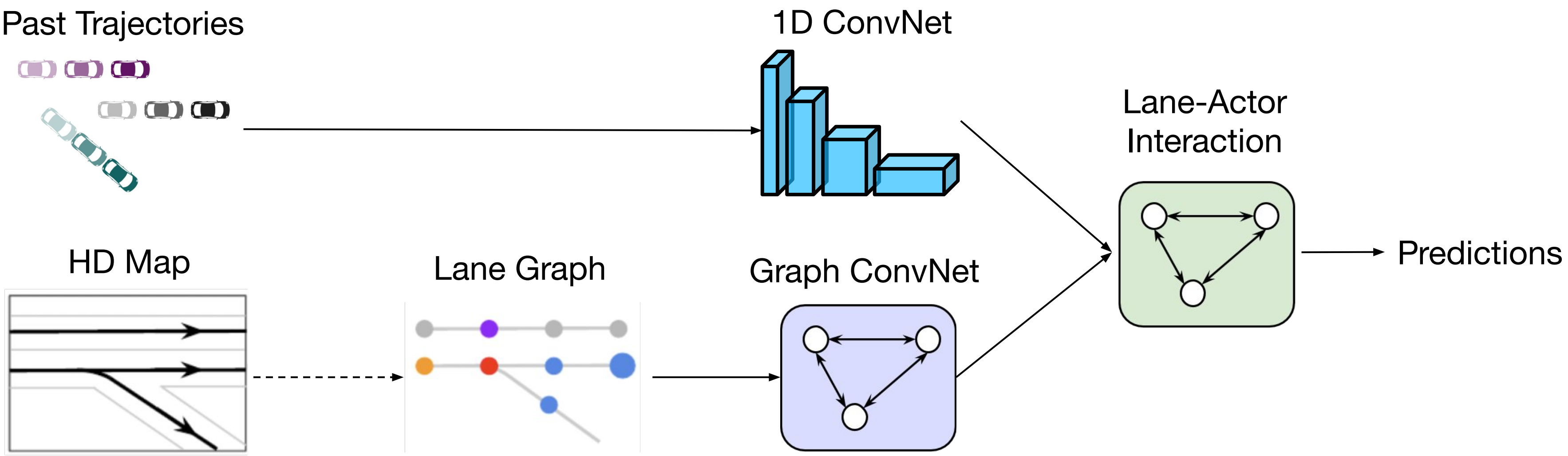
- Lossy rendering of both trajectories and HD map
- 2D convolution on raster images is computation-intensive



[1] Short-term Motion Prediction of Traffic Actors for Autonomous Driving using Deep Convolutional Networks. [N. Djuric, et al. 2018]
[2] ChauffeurNet: Learning to Drive by Imitating the Best and Synthesizing the Worst. [M. Bansal, et al. 2018]

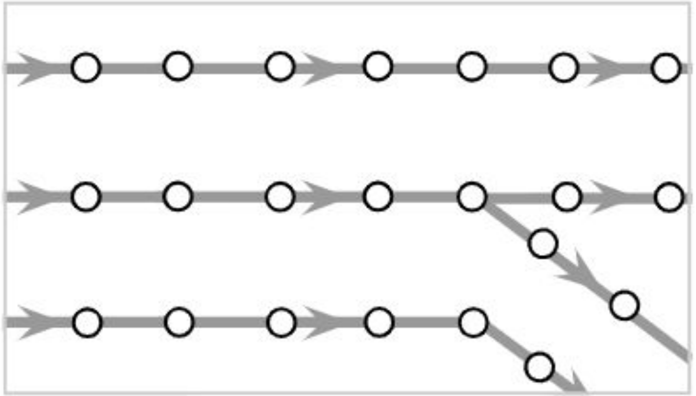
Our Approach: Lane Graph

- Minimal information loss of map geometry and semantics
- Efficient and effective feature learning on graph-structured data

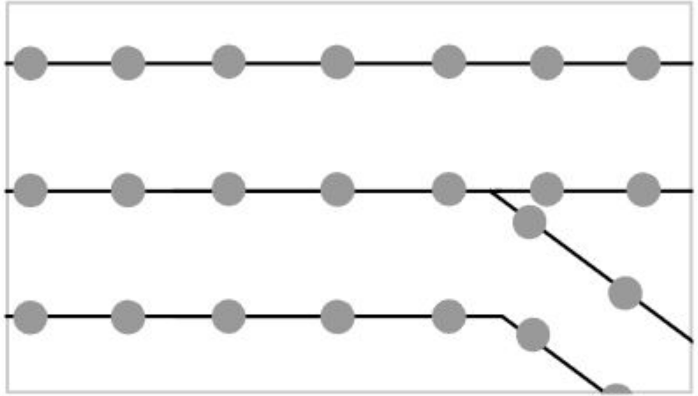


Lane Graph: Nodes

Raw map data



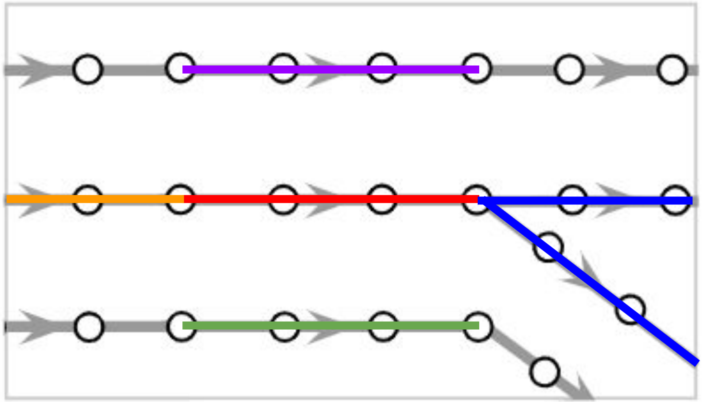
Our lane graph



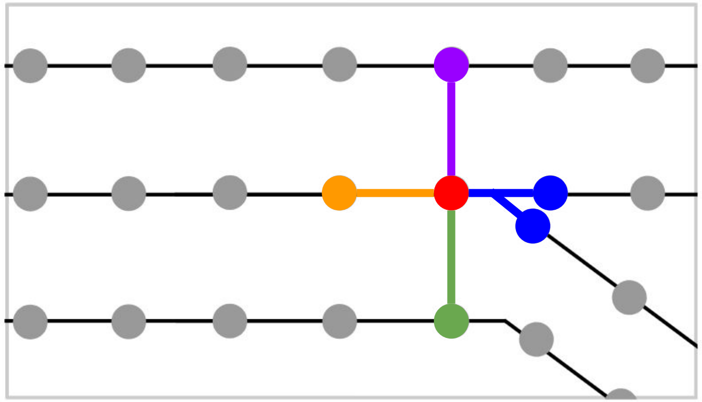
- Raw map:
    - A set of *directed polylines* representing the lane centerlines
- Lane graph:
    - Each node represents one *directed line segment*
    - Preserves full geometric shape, enables fine-grained lane-actor interaction

Lane Graph: Edges

Raw map data

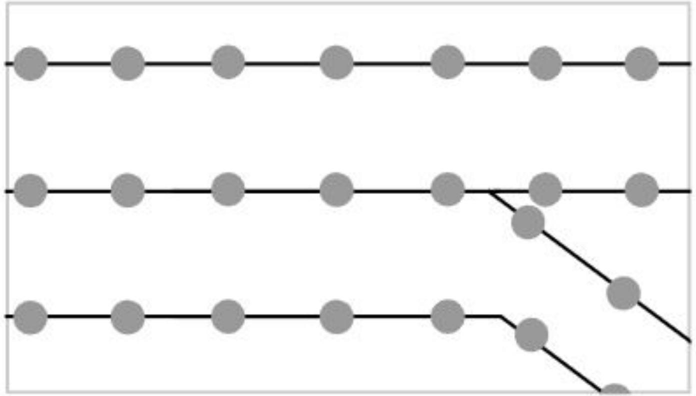


Our lane graph



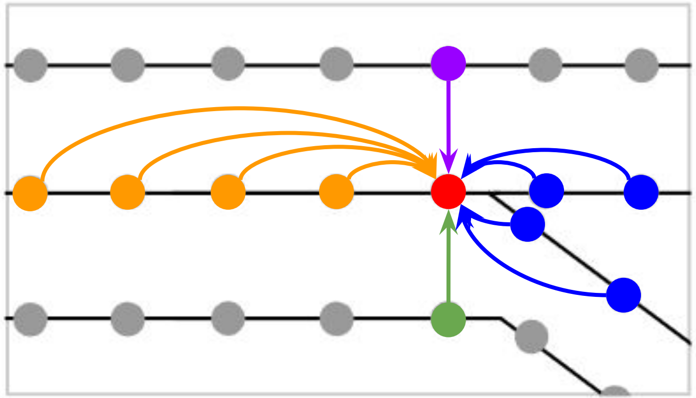
- Raw map:
  - 4 connectivity types: predecessor, successor, left neighbor, right neighbor
- Lane graph:
  - Multi-type & sparse connectivity between nodes
  - Enables structured information propagation

Lane Graph: Node Feature



- Node feature initialization: $\mathbf{x}_i = \mathrm{MLP}_{\mathrm{shape}}\left(\mathbf{v}_i^{\mathrm{end}} - \mathbf{v}_i^{\mathrm{start}}\right) + \mathrm{MLP}_{\mathrm{loc}}\left(\mathbf{v}_i\right)$

Lane Graph: Node Feature Update

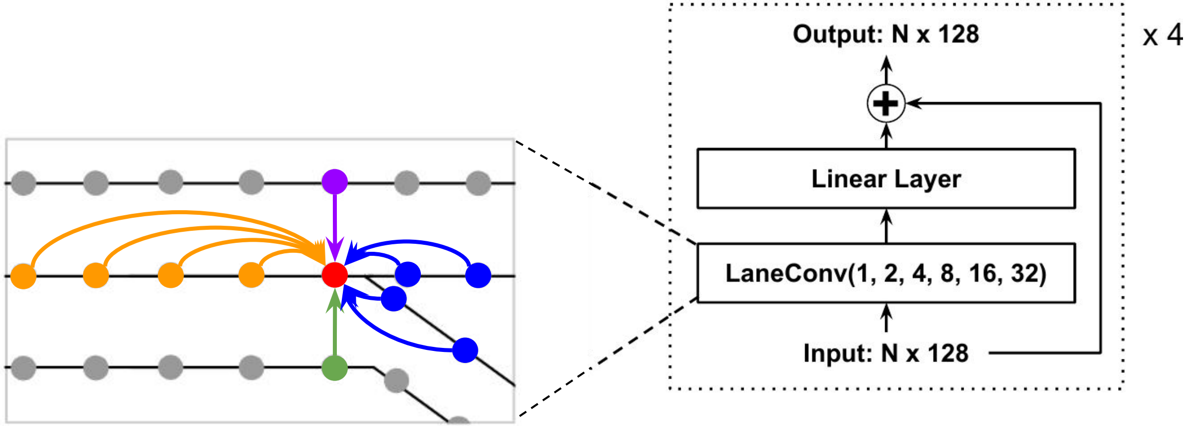


- Multi-scale LaneConv: $Y = XW_0$ Self

$$+ \sum_{i \in \{\text{left},\text{right}\}} A_i X W_i$$ Left neighbors & right neighbors

$$+ \sum_{c=1}^{C} \left(A_{\text{pre}}^{k_c} X W_{\text{pre},k_c} + A_{\text{suc}}^{k_c} X W_{\text{suc},k_c} \right)$$ *Multi-scale* predecessors & successors

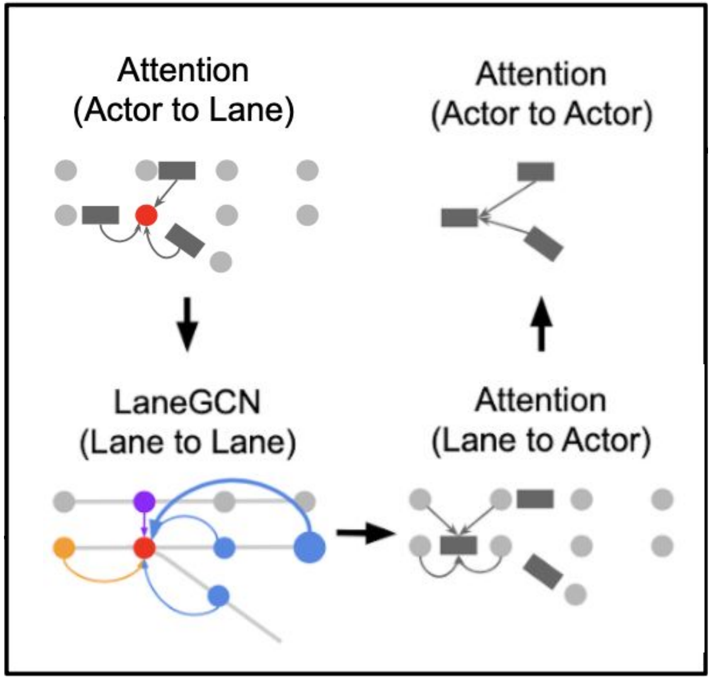
LaneGCN: Network Architecture



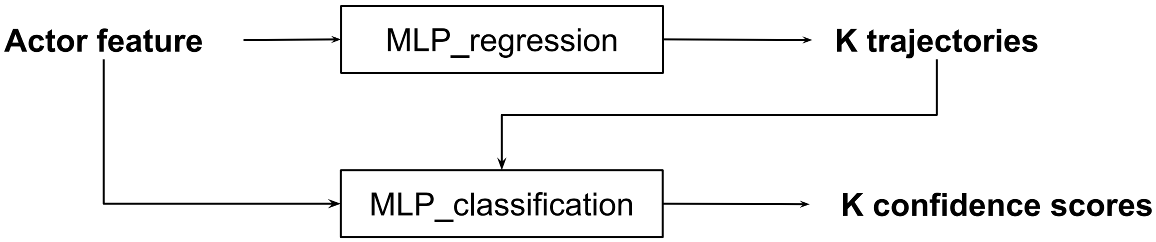
- We apply a variant of graph convnet (namely LaneGCN) on the lane graph to extract node features
- LaneGCN architecture: a stack of 4 multi-scale LaneConv blocks

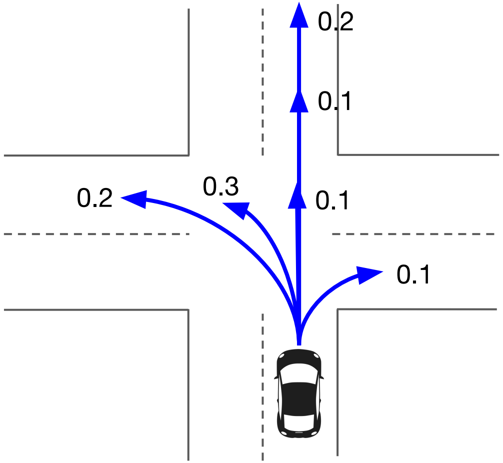
4-Way Lane-Actor Interactions

- **Actor-to-Lane**: Propagate real-time traffic information to lane features. For example, if a lane is occupied.
- **Lane-to-Lane**: Propagate the traffic information along the lane graph.
- **Lane-to-Actor**: Fuse the latest lane information back to actors.
- **Actor-to-Actor**: Interaction between actors.



Prediction Header





- Input: actor feature after 4-way lane-actor interactions
- Two branch outputs:
  - Regression: output K future trajectories
  - Classification: output K confidence scores conditioned on both actor feature and predicted trajectories

Evaluation Results on Argoverse

Model	K=1			K=6		
	minADE	minFDE	MR	minADE	minFDE	MR
Argoverse Baseline	2.96	6.81	0.81	2.34	5.44	0.69
Argoverse Baseline (NN)	3.45	7.88	0.87	1.71	3.29	0.54
Holmes (*7th*)	2.91	6.54	0.82	1.38	2.66	0.42
cxx (*3rd*)	1.91	4.31	0.66	0.99	1.71	0.19
uulm-mrm (*2nd*)	1.90	4.19	0.63	0.94	1.55	0.22
Jean (*1st*)	1.86	4.18	0.63	0.93	1.49	0.19
Our Model	**1.71**	**3.78**	**0.59**	**0.87**	**1.36**	**0.16**

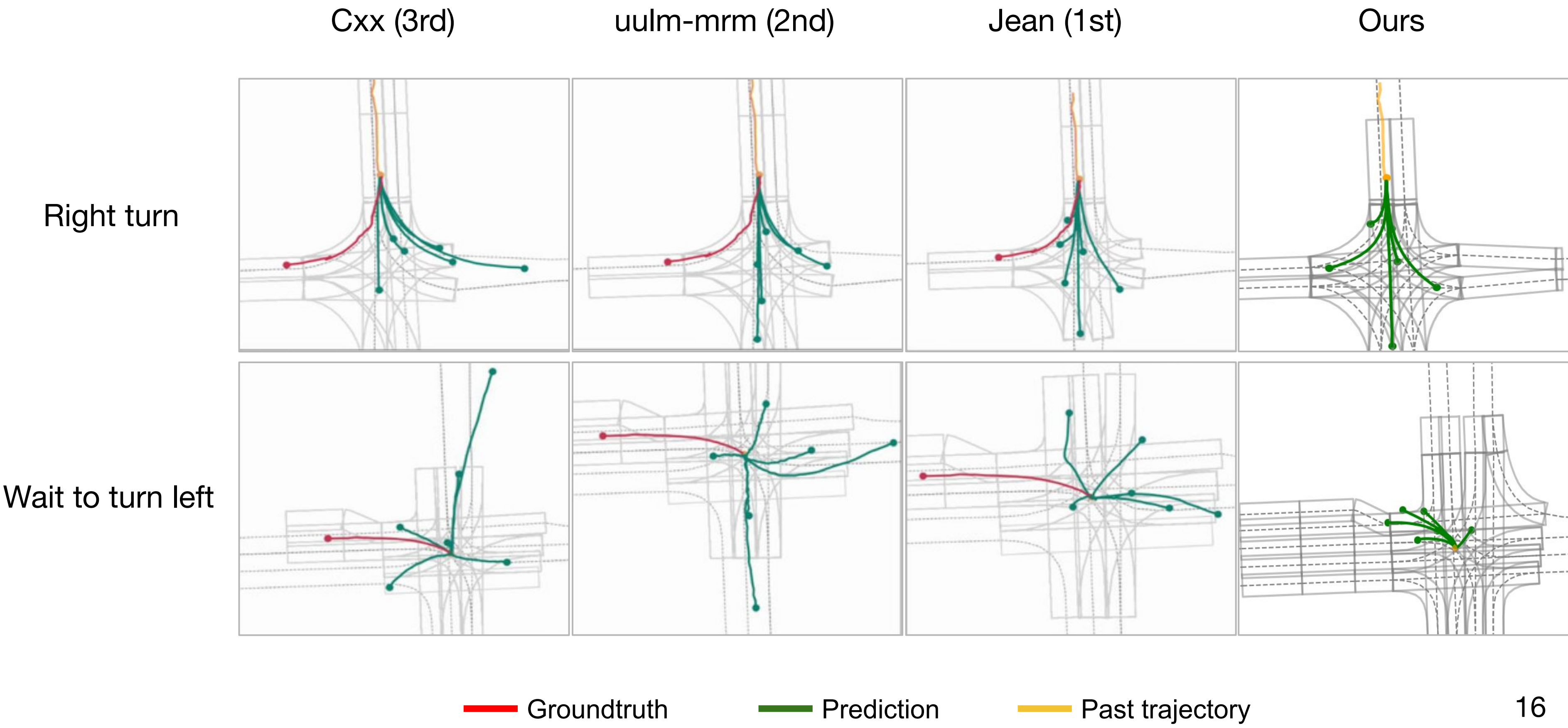
Ablation Study on Modules

Backbone		Fusion Cycle				K=1		K=6	
ActorNet	MapNet	A2L	L2L	L2A	A2A	minADE	minFDE	minADE	minFDE
✓						1.90	4.38	0.91	1.66
✓					✓	1.58	3.61	0.79	1.29
✓	✓			✓		1.55	3.52	0.76	1.23
✓	✓	✓	✓	✓	✓	**1.35**	**2.97**	**0.71**	**1.08**

Ablation Study on Graph Operators

Component				K=1		K=6	
GraphConv	Residual	Multi-Type	Dilate	minADE	minFDE	minADE	minFDE
✓				1.72	3.93	0.82	1.41
✓	✓			1.53	3.48	0.79	1.33
✓	✓	✓		1.48	3.33	0.74	1.19
✓	✓	✓	✓	**1.39**	**3.05**	**0.72**	**1.10**

Qualitative Comparison on Argoverse

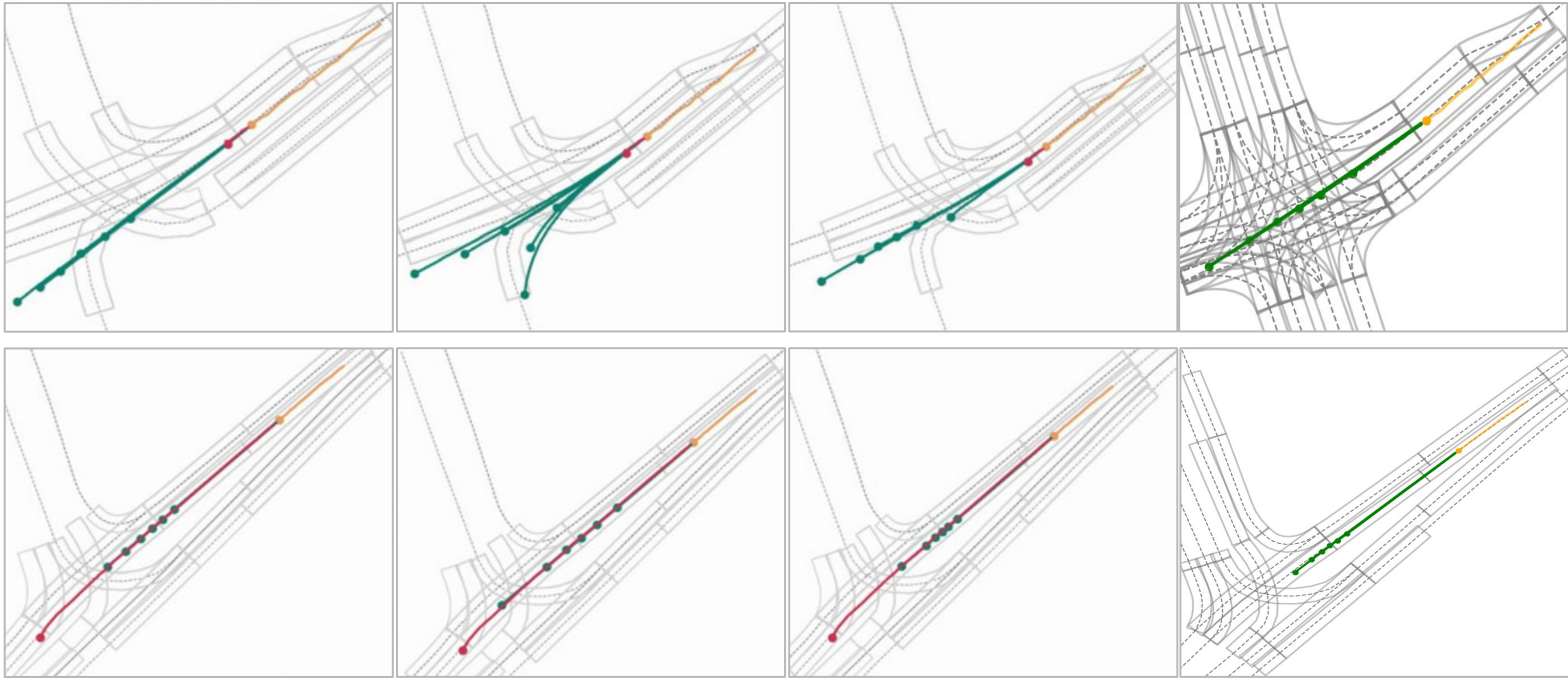


Qualitative Comparison on Argoverse

Cxx (3rd) | uulm-mrm (2nd) | Jean (1st) | Ours

Deceleration

Acceleration



Groundtruth Prediction Past trajectory

Conclusion

- A new representation for HD maps: lane graph
- A new operator for feature extraction on lane graph: multi-scale LaneConv
- 4-way interactions between lanes and actors
- New state-of-the-art results on the Argoverse benchmark

Learning Lane Graph Representations for Motion Forecasting. [M. Liang, et al. ECCV 2020]