Görkay Aydemir,   Adil Kaan Akan,   Fatma Güney

KUIS AI Center,   Koç University

Multi-Agent Trajectory Prediction

Goal: Predict future locations of **all** agents in the scene

Motivation: For safe navigation of autonomous vehicles, possible future behaviors of surrounding vehicles should be known. Considering real-world requirements, the trade-off in future prediction between accuracy and efficiency should be solved. We benefit from advantages of both scene-centric and agent-centric approaches for efficient and accurate multi-agent trajectory prediction

(a) Scene-Centric

(b) Agent-Centric

In scene-centric approach (a), all elements are encoded according to the same reference point. In agent-centric approach (b), each agent is encoded in its own reference frame, leading to a complexity linear in the number of agents

Approach:

- Changing reference point from the agent to the scene i.e.agent-centric to **scene-centric**, providing efficiency
- Using **dynamic weight learning** to adapt to the situation of each agent, enabling us to predict trajectories of multiple agents accurately
- Utilizing a **goal-conditioned architecture** to reach a unified prediction process for both single and multi-agent settings
- Stabilizing prediction with **gradient stopping**, achieving SOTA results using a small and fast model both in single-agent and multi-agent prediction

Methodology

| Input Vectors | Features | Endpoints | Offsets | // Gradient Detach |

Feature Encoding: Encode all items, represented in vectorized format, with transformers

Endpoint Prediction: Predict and refine possible endpoints using adaptive head

Trajectory Prediction: Interpolate middle steps conditioned on endpoints

Results

(a) Multi-Agent

(b) Single-Agent

We visualize multi-agent predictions on Interaction and single-agent on Argoverse

Accuracy vs. Efficiency on Argoverse. Our method achieves the lowest error with a small number of parameters, leading to highly efficient inference time compared to methods AutoBot, LaneGCN, mmTransformer, DenseTNT, SceneTransformer, LTP, PAGA, and HiVT.

	mADE$_k$		mFDE$_k$		brier-	#Prm
	k=1	k=6	k=1	k=6	mFDE$_6$	(M)
THOMAS	1.67	0.94	3.59	1.44	1.897	-
SceneTr	1.81	0.80	4.05	1.23	1.89	15.3
LTP	1.62	0.83	3.55	1.30	1.86	**1.1**
HiVT	1.60	**0.77**	3.53	**1.17**	1.84	2.5
PAGA	**1.56**	0.80	**3.38**	1.21	**1.76**	1.6
Ours (Single)	1.59	0.79	3.50	**1.17**	1.80	1.4

Results of single-agent prediction on Argoverse

	mADE$_6$	mFDE$_6$	MR$_6$	Inf. (ms)
AutoBot (J.)	0.21	0.64	0.06	25.29
SceneTr	0.26	0.47	0.05	-
THOMAS	0.26	0.46	0.05	-
Ours (Multi)	**0.16**	**0.34**	**0.01**	11.10

Results of multi-agent prediction on Interaction

Contact Information

Website: kuis-ai.github.io/adapt/

E-mail: {gaydemir23, kakan20, fguney}@ku.edu.tr