

Planning-oriented Autonomous Driving

Yihan Hu* Jiazhi Yang* Li Chen*[†] Keyu Li*

Chonghao Sima Xizhou Zhu Siqi Chai Senyao Du Tianwei Lin Wenhui Wang
Lewei Lu Xiaosong Jia Qiang Liu Jifeng Dai Yu Qiao Hongyang Li[†]

*equal contribution [†]project lead



Yihan



Jiazhi



Li



Keyu



Hongyang



Poster: THU-AM-131

arXiv: <https://arxiv.org/abs/2212.10156>



上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

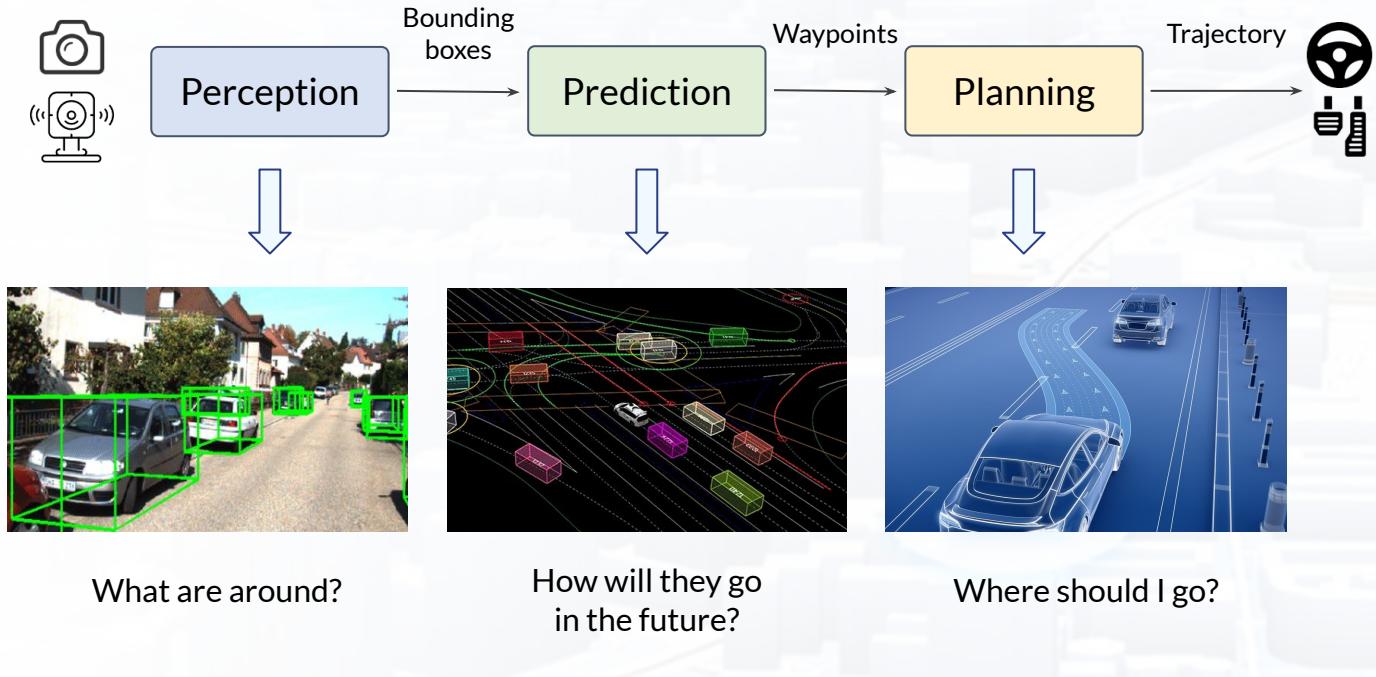
Shanghai AI Laboratory | 上海人工智能实验室



Planning-oriented Autonomous Driving

Background and Motivation

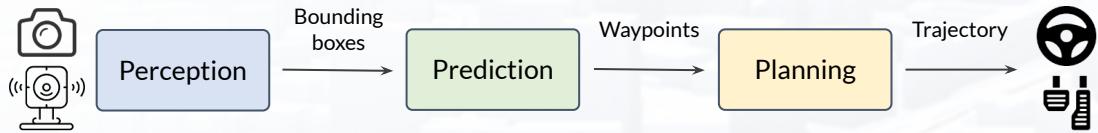
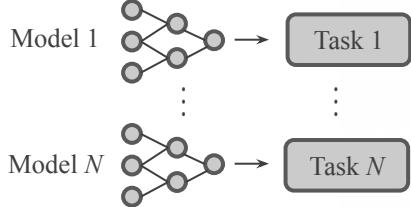
Background - Autonomous Driving (AD) Systems



Various weathers, illuminations,
and scenarios

Background - Design Options for Autonomous Driving (AD) Systems

(a) Standalone Models



- Typical Industry solutions
- Independent teams for module developments
- ✗ • Severe error accumulation

Isolated Optimization Objective



Object Detection



Motion Prediction



Planning

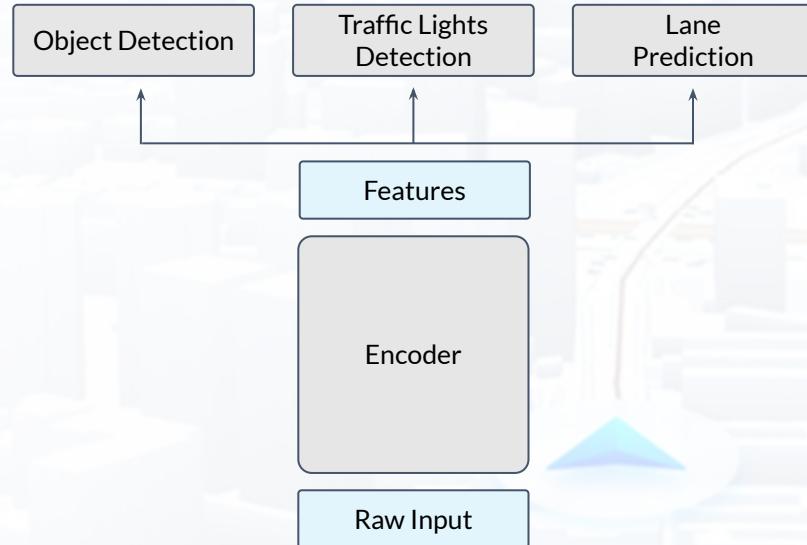
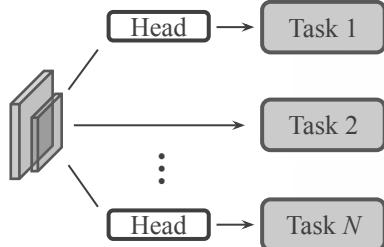
Optimization metric **mAP**

Optimization metric **minFDE**

Optimization target **Safety and Comfort**

Background - Design Options for Autonomous Driving (AD) Systems

(b) Multi-task Framework



- Shared feature for multiple tasks
- Easily extended to more tasks,
Compute-efficient
- ✗ • Lack of tasks' coordination

credit to Tesla AI Day 2021

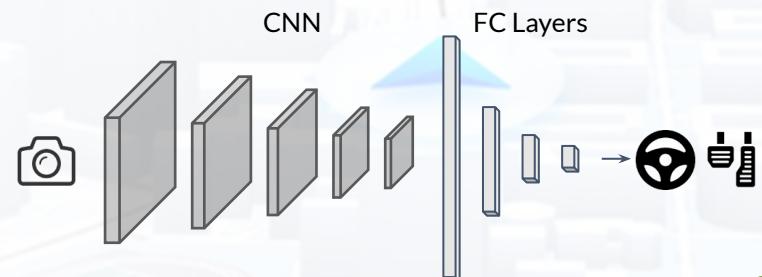
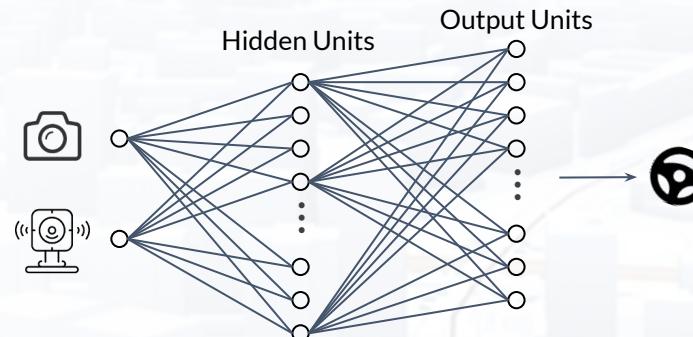


Background - Design Options for Autonomous Driving (AD) Systems

(c.1) End-to-end Framework - Vanilla Solutions



- Direct policy learning from sensor inputs, bypassing intermediate tasks
- Simple design with good performance in the simulator
- ✗ • Deficient in interpretability

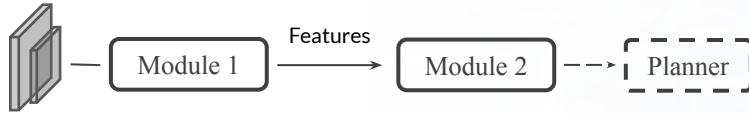


DAVE-2, arXiv 2016. Nvidia



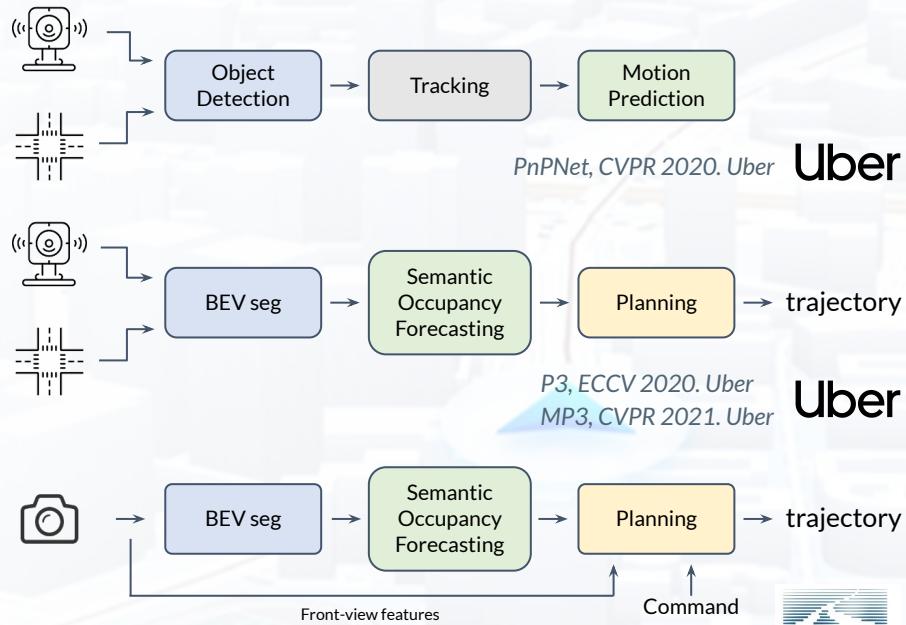
Background - Design Options for Autonomous Driving (AD) Systems

(c.2) End-to-end Framework - Explicit / Interpretable Design



- Introducing **intermediate tasks** to assist planning
- Better interpretability (e.g. Bird's-eye-view, BEV)
- ✗ • Lack some crucial components¹

1. *The necessities of each component is mentioned in Appendix.*

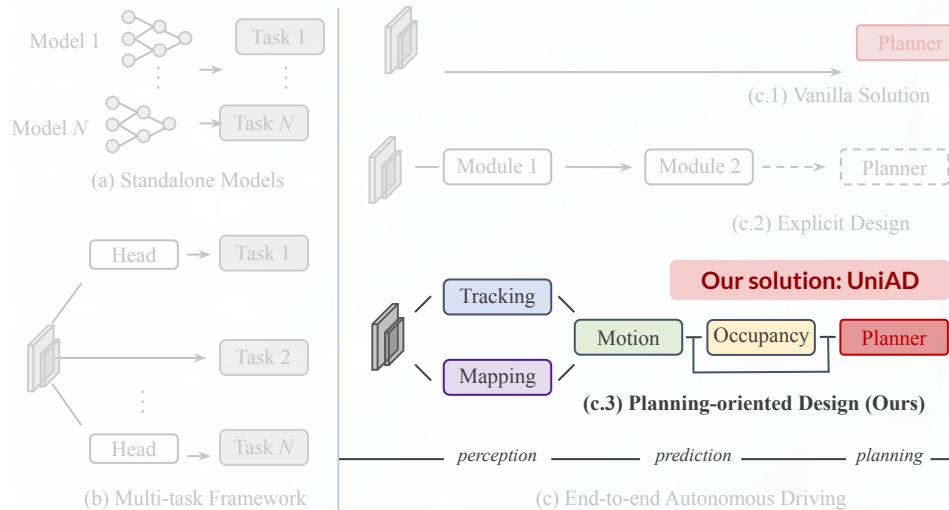


ST-P3, ECCV 2022, SH AI Lab



Motivation- Towards Reliable Planning

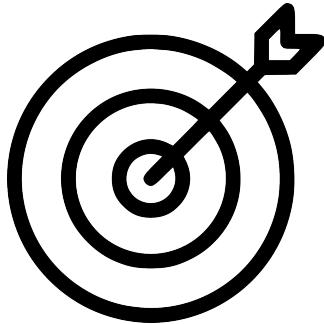
Ours: Planning-oriented Autonomous Driving



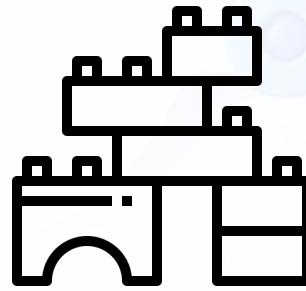
What do we want:

- ✓ • **Unify full-stack AD tasks**
- ✓ • **Coordinate all task towards safe planning**

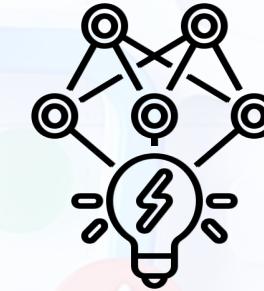
UniAD - Overview



Which tasks?



How to construct?



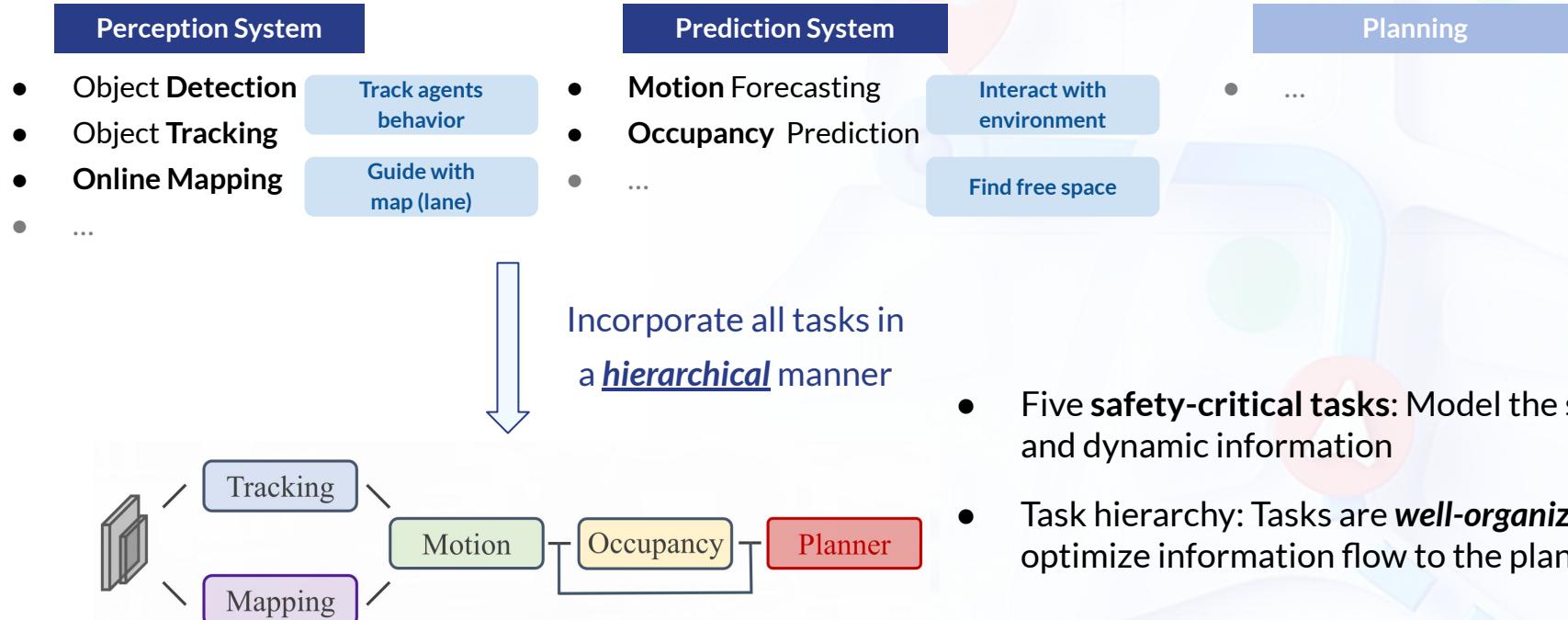
How to train?



Planning-oriented Autonomous Driving

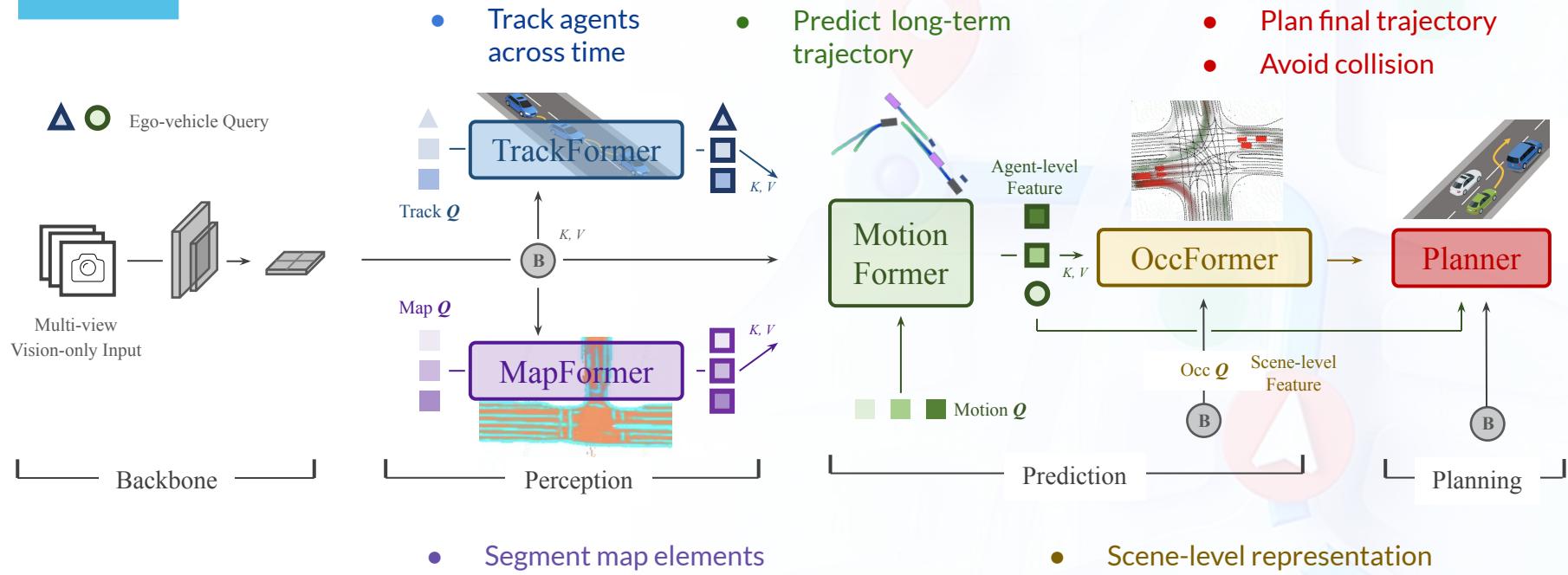
Delving into Details

UniAD - Which Tasks?



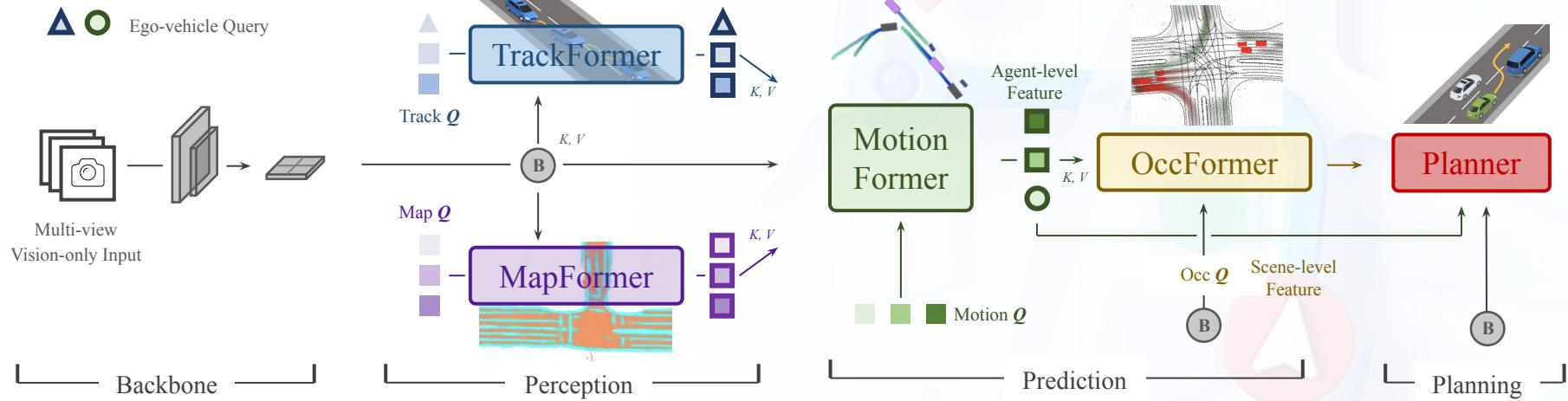
UniAD - How to Construct?

Pipeline



UniAD - How to Construct?

Pipeline



- Entire pipeline connected by queries
- Tasks coordinated with queries
- Interactions modeled by attention

Unified Query

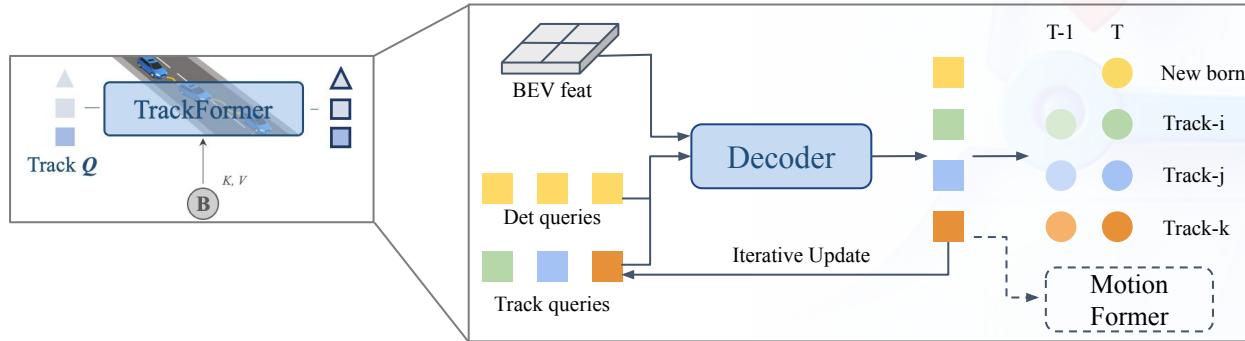
Transformer-based

First time to unify
full-stack AD tasks!

UniAD - How to Construct?

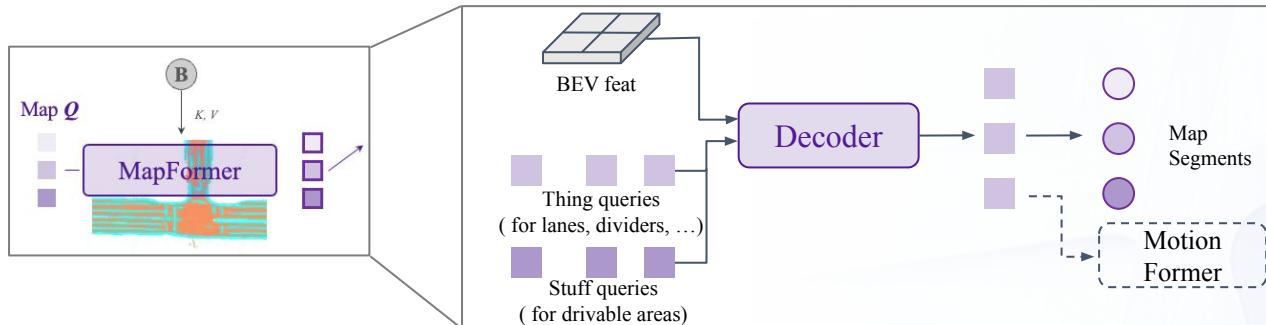
Perception → Prediction → Planning

TrackFormer - MOTR (ECCV 2022)



- End-to-end trainable tracking without post-association

MapFormer - Panoptic SegFormer (CVPR 2022)



- Each query represents a map element

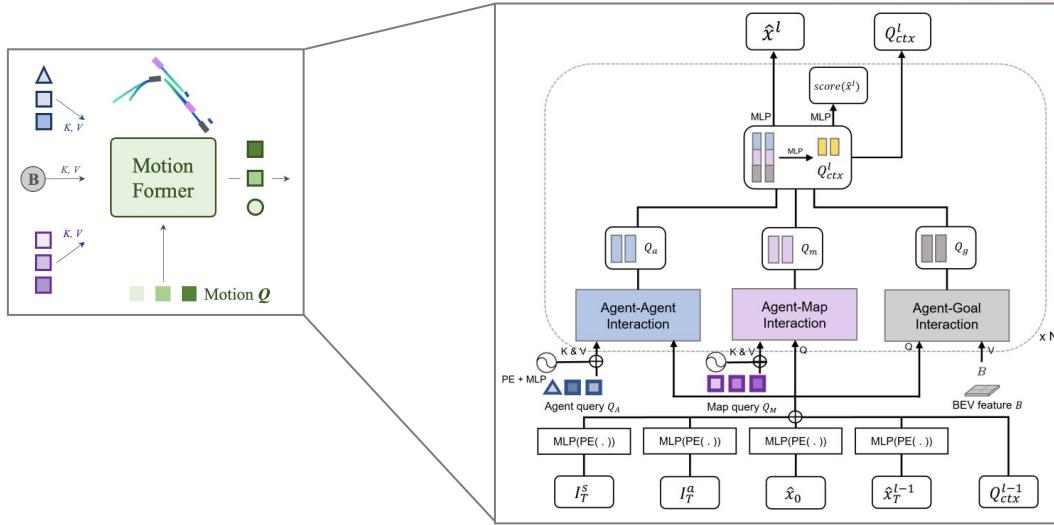
UniAD - How to Construct?

Perception

Prediction

Planning

MotionFormer (Proposed in UniAD)



- Diverse **relation modelings** via attentions:
Agent-agent, agent-map, agent-goal

- Non-linear optimization:**
Adjust ground-truth trajectory
based on upstream predictions



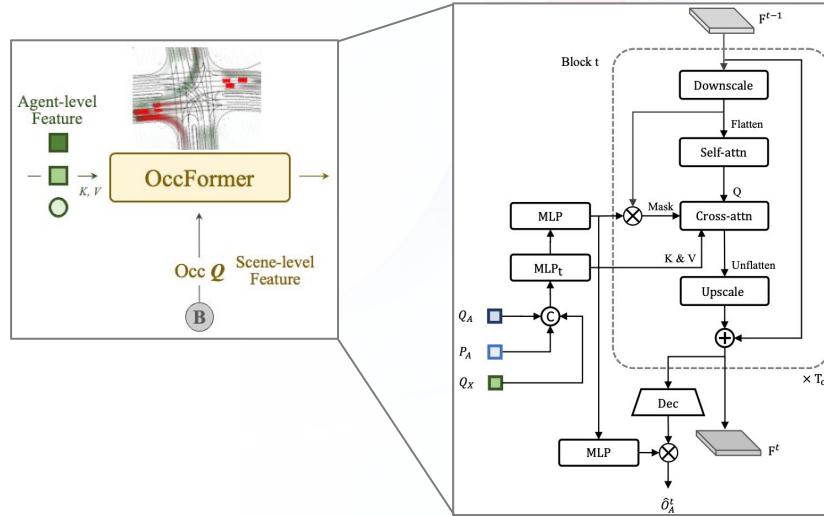
UniAD - How to Construct?

Perception

Prediction

Planning

OccFormer (Proposed in UniAD)



- **Encode agent-wise knowledge into the scene representation**
- **Predict occupancy as attention mask to restrict the interactions between the agents and their corresponding BEV features.**

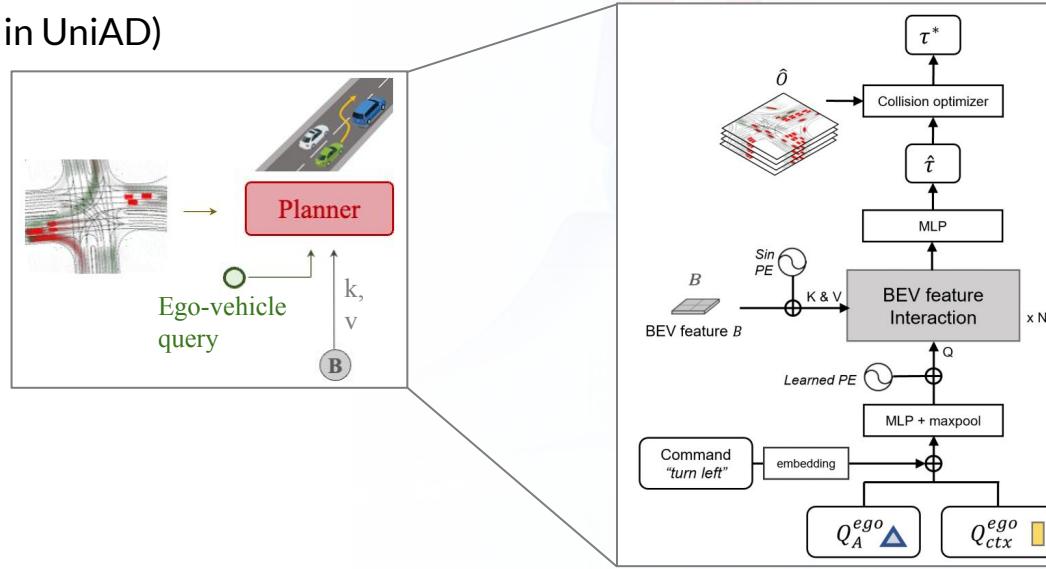
UniAD - How to Construct?

Perception

Prediction

Planning

Planner (Proposed in UniAD)

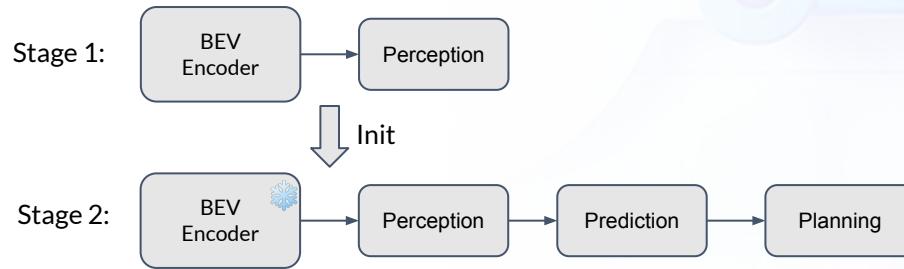


- **Ego-vehicle query:** consistently models the ego-vehicle
- **Collision optimization:** Steer the predicted trajectories clear of predicted occupancy.

The Recipe - How to Train?

Two-phase training. Perception stage + End-to-end stage

- The stabilized perception capability helps the end-to-end stage **converge faster**



Shared matching. Matching results of tracking reused in motion and occupancy

- Consistent learning of agent identities
- Converging faster



Planning-oriented Autonomous Driving Experiments

UniAD - Ablation Results

Tasks benefit  each other and contribute to safe planning

ID	Modules					Tracking			Mapping		Motion Forecasting			Occupancy Prediction			Planning		
	Track	Map	Motion	Occ.	Plan	AMOTA↑	AMOTP↓	IDS↓	IoU-lane↑	IoU-road↑	minADE↓	minFDE↓	MR↓	IoU-n.↑	IoU-f.↑	VPQ-n.↑	VPQ-f.↑	avg.L2↓	avg.Col.↓
0*	✓	✓	✓	✓	✓	0.356	1.328	893	0.302	0.675	0.858	1.270	0.186	55.9	34.6	47.8	26.4	1.154	0.941
1	✓					0.348	1.333	791	-	-	-	-	-	-	-	-	-	-	
2		✓				-	-	-	0.305	0.674	-	-	-	-	-	-	-	-	
3	✓	✓				0.355	1.336	785	0.301	0.671	-	-	-	-	-	-	-	-	
4			✓			-	-	-	-	-	0.815	1.224	0.182	-	-	-	-	-	
5	✓		✓			<u>0.360</u>	1.350	919	-	-	0.751	1.109	0.162	-	-	-	-	-	
6	✓	✓	✓			0.354	1.339	820	0.303	0.672	0.736(-9.7%)	1.066(-12.9%)	0.158	-	-	-	-	-	
7				✓		-	-	-	-	-	-	-	-	60.5	37.0	52.4	29.8	-	
8	✓			✓		<u>0.360</u>	1.322	809	-	-	-	-	-	62.1	38.4	52.2	32.1	-	
9	✓	✓	✓	✓		0.359	1.359	1057	0.304	0.675	0.710(-3.5%)	1.005(-5.8%)	0.146	62.3	39.4	53.1	32.2	-	-
10					✓	-	-	-	-	-	-	-	-	-	-	-	1.131	0.773	
11	✓	✓	✓		✓	0.366	1.337	889	0.303	0.672	0.741	1.077	0.157	-	-	-	-	1.014	0.717
12	✓	✓	✓	✓	✓	0.358	<u>1.334</u>	641	0.302	0.672	<u>0.728</u>	<u>1.054</u>	<u>0.154</u>	62.3	39.5	52.8	32.3	1.004	0.430

Conclusion:

- ID. 4-6: Track & Map → Motion 
- ID. 7-9: Motion  ↔ Occupancy 
- ID. 10-12: Motion & Occupancy → Planning 

UniAD - Results

Even outperforms LiDAR-based counterparts on planning

†: LiDAR-based
Camera-based

Planning

Method	L2($m\downarrow$)				Col. Rate(%) \downarrow			
	1s	2s	3s	Avg.	1s	2s	3s	Avg.
NMP [†] [88]	-	-	2.31	-	-	-	1.92	-
SA-NMP [†] [88]	-	-	2.05	-	-	-	1.59	-
FF [†] [36]	0.55	1.20	2.54	1.43	0.06	0.17	1.07	0.43
EO [†] [42]	0.67	1.36	2.78	1.60	0.04	0.09	0.88	0.33
ST-P3 [37]	1.33	2.11	2.90	2.11	0.23	0.62	1.27	0.71
UniAD	0.48	0.96	1.65	1.03	0.05	0.17	0.71	0.31

UniAD - Results

SOTA performance on all investigated tasks

Multi-object Tracking

Method	AMOTA↑	AMOTP↓	Recall↑	IDS↓
Immortal Tracker [†] [82]	0.378	1.119	0.478	936
ViP3D [30]	0.217	1.625	0.363	-
QD3DT [35]	0.242	1.518	0.399	-
MUTR3D [91]	0.294	1.498	0.427	3822
UniAD	0.359	1.320	0.467	906

Mapping

Method	Lanes↑	Driveable↑	Divider↑	Crossing↑
VPN [63]	18.0	76.0	-	-
LSS [66]	18.3	73.9	-	-
BEVFormer [48]	23.9	77.5	-	-
BEVerse [†] [92]	-	-	30.6	17.2
UniAD	31.3	69.1	25.7	13.8

Motion Forecasting

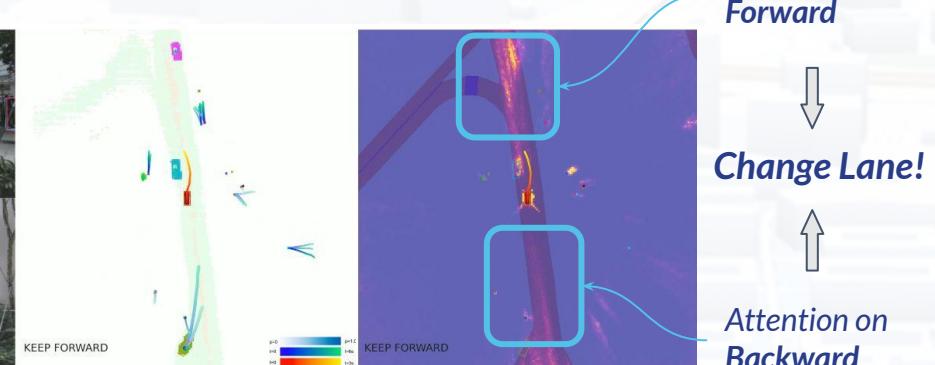
Method	minADE(m)↓	minFDE(m)↓	MR↓	EPA↑
PnPNet [†] [50]	1.15	1.95	0.226	0.222
ViP3D [30]	2.05	2.84	0.246	0.226
Constant Pos.	5.80	10.27	0.347	-
Constant Vel.	2.13	4.01	0.318	-
UniAD	0.71	1.02	0.151	0.456

Occupancy Prediction

Method	IoU-n.↑	IoU-f.↑	VPQ-n.↑	VPQ-f.↑
FIERY [34]	59.4	36.7	50.2	29.9
StretchBEV [1]	55.5	37.1	46.0	29.0
ST-P3 [37]	-	38.9	-	32.1
BEVerse [†] [92]	61.4	40.9	54.3	36.1
UniAD	63.4	40.2	54.7	33.5

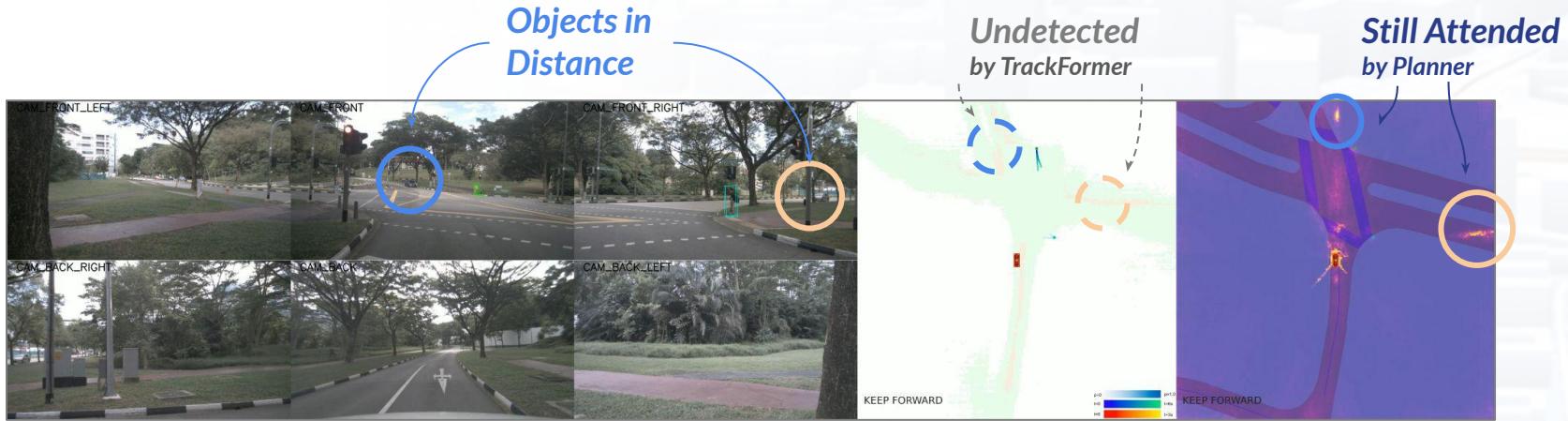
UniAD - Visualizations

Planner attends to crucial areas in complex scenes



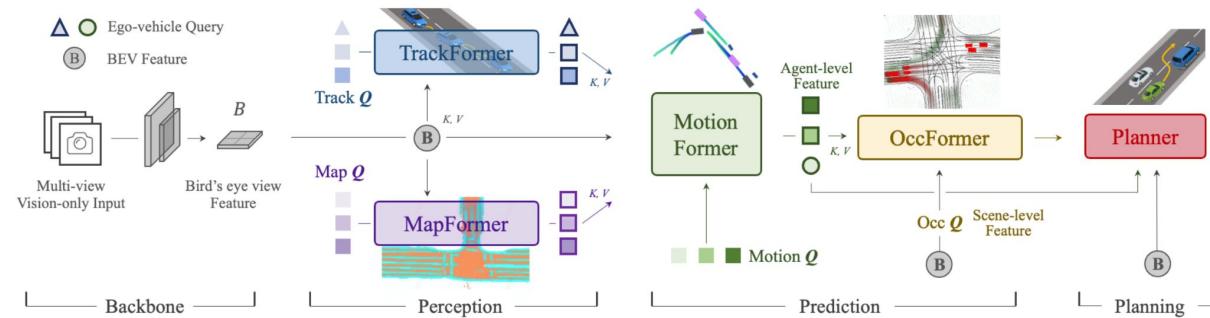
UniAD - Recover from Upstream Errors

Planner could still attend to ‘undetected’ regions/objects



One-page Summary

- **Planning-oriented Philosophy:** An end-to-end autonomous driving (AD) framework in pursuit of safe planning, equipped with a wide span of AD tasks.
- **Unified Query design:** Queries as interfaces to connect and coordinate all tasks.
- **State-of-the-art (SOTA) Performance** with vision-only input.
- **First Step towards Autonomous Driving Foundation Models**





What's next? beyond UniAD

Embracing Foundation Models for Autonomous Driving



Data & Training Strategy

- Multiple datasets with labels for various tasks?

Shippable Algorithm

- More modules integration, extensible to applications (e.g. V2X)

Closed-loop System

- Closed-loop training and testing in simulator & real world

Check out the latest **Survey Paper**!

[https://github.com/OpenDriveLab/
End-to-end-Autonomous-Driving](https://github.com/OpenDriveLab/End-to-end-Autonomous-Driving)



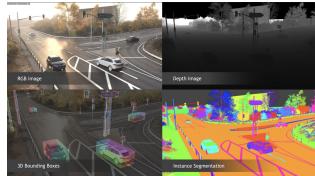
Beyond UniAD: DriveAGI

Data-centric Pipeline

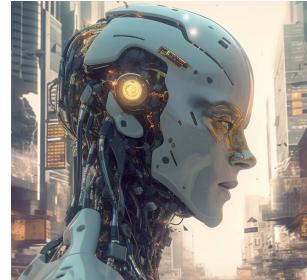
Data Collection



Data Generation



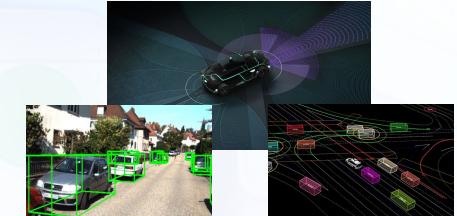
Pre-training DriveCore



How to formulate?
What's the objective goal?

Applications

Autonomous Driving



Broader Impact



Partial photo by courtesy of online resources.

OpenDriveLab



Poster: THU-AM-131

THANKS

<https://opendrivelab.com>



上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

