Divide and Merge:
Motion and Semantic Learning in End-to-End Autonomous Driving

Yinzhe Shen [1] Ömer Şahin Taş [1,2] Kaiwen Wang [1] Royden Wagner [1] Christoph Stiller [1,2]

[1]Karlsruhe Institute of Technology (KIT) [2]FZI Research Center for Information Technology

Karlsruhe Institute of Technology

TLDR: solving the perception negative transfer problem boosts driving safety

The Problem: Current E2E models use a single feature vector to represent both **semantics** (what is it?) and **motion** (where is it going?). Forcing features to learn motion (prediction/planning) impairs their ability to represent semantics (detection/tracking). Perception performance drops when jointly trained with prediction and planning, which is known as **perception negative transfer**.

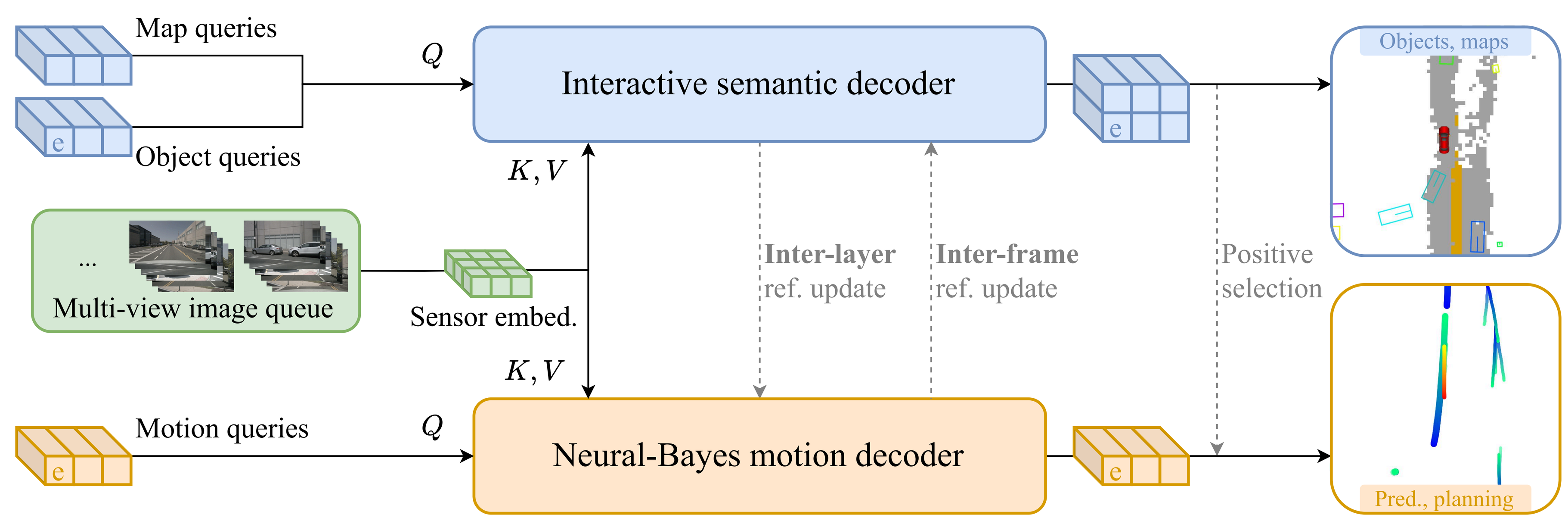
Our Solution (DMAD):

- **Divide:** Decouple heterogeneous tasks (motion vs. semantics) into parallel pathways, **mitigating negative** transfer.
- **Merge:** Enhance similar tasks (object vs. map) via interactive attention, **promoting positive** transfer.

Results: performance improvements across **all tasks** (object & map perception, prediction, and planning).

3-min video

Overview: the DMAD framework



Structure: Parallel semantic and motion learning in two pathways.

1. Divide: Neural-Bayes motion decoder

Goal: Decouple motion from semantics to stop negative transfer.

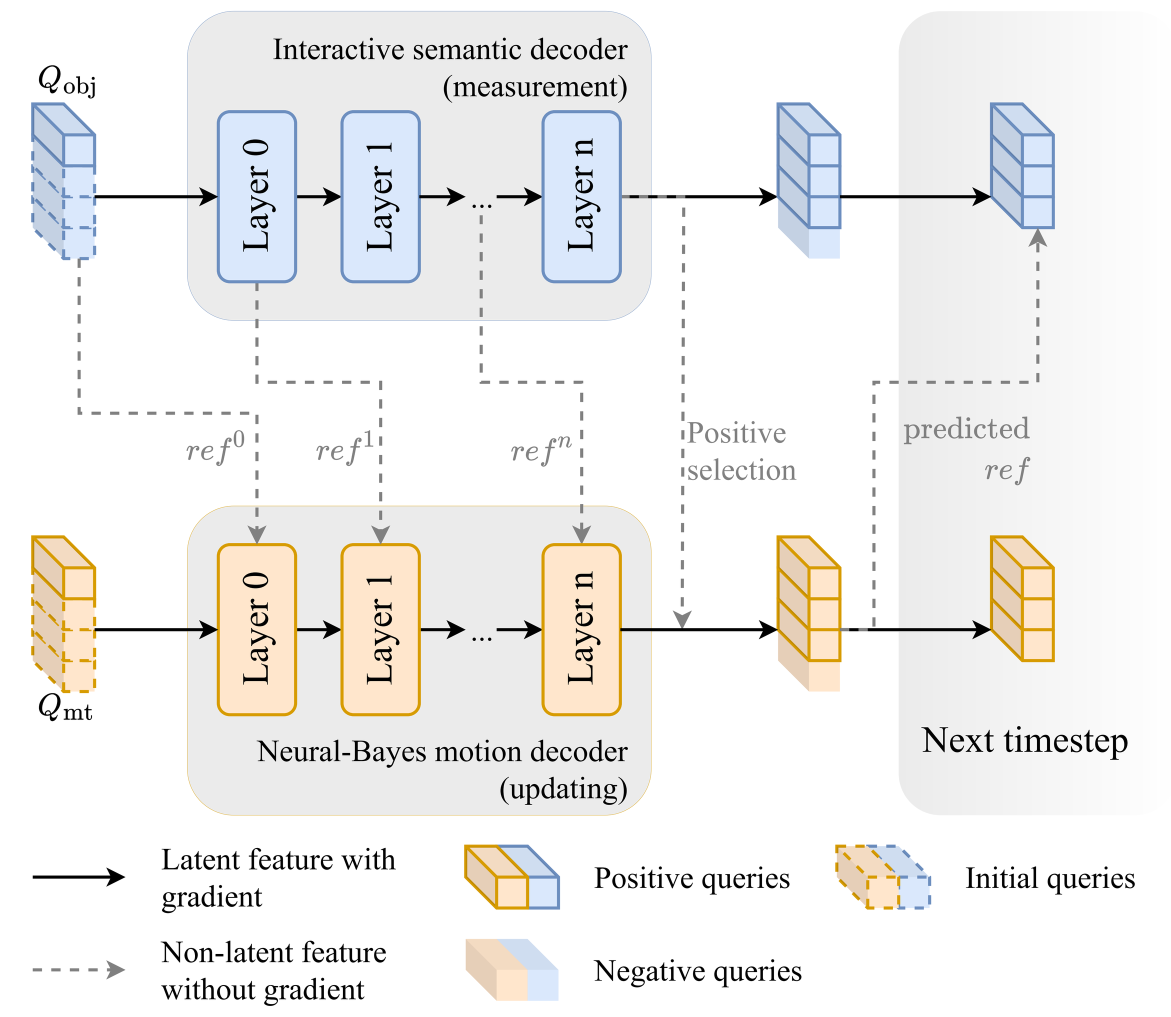
- **Object representation:** A query pair (Q_{mt} and Q_{obj}) represents an object instance.
- **Bayes filter inspiration:** Measurement, updating and prediction.
- **Recursive design:** Recursively exchanging reference points between both kinds of queries.

2. Merge: interactive semantic decoder

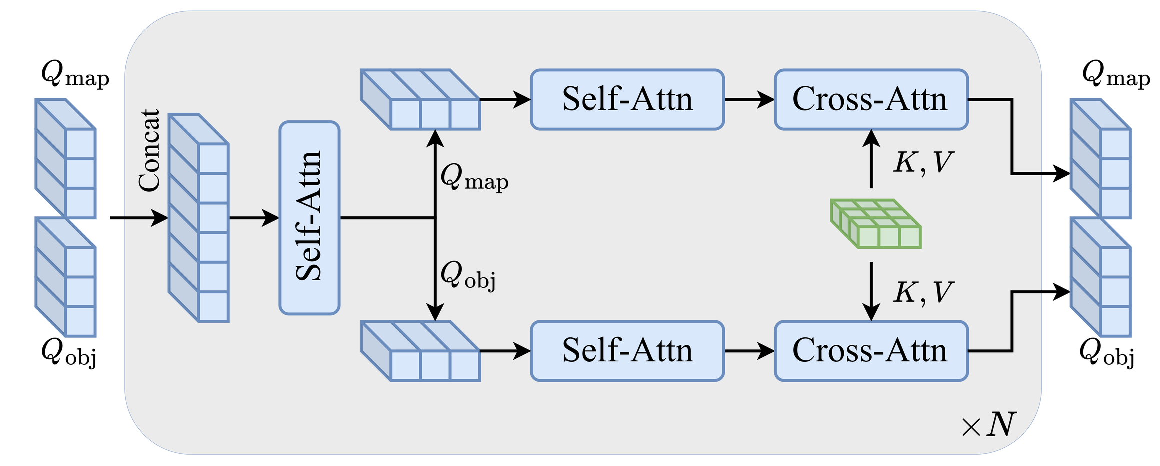
Goal: Enhance semantic consistency between objects and maps.

- **Intuition:** Cars are likely to be on drivable areas.
- **Mechanism:** A self-attention module allows Q_{obj} and Q_{map} to exchange context.

Module diagrams



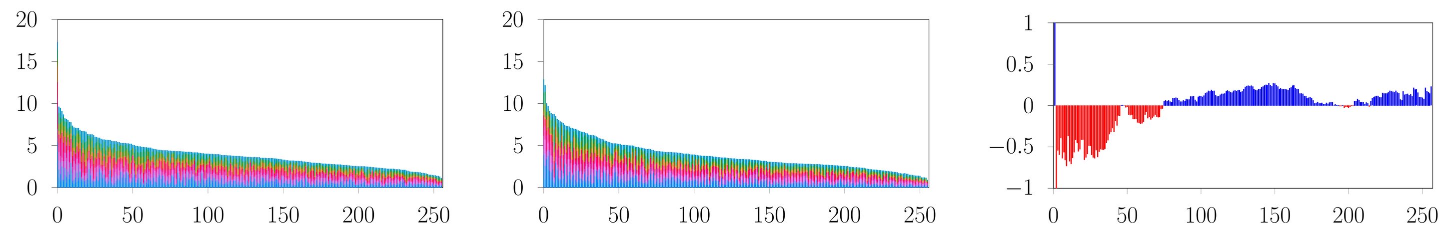
Divide: Neural-Bayes motion decoder



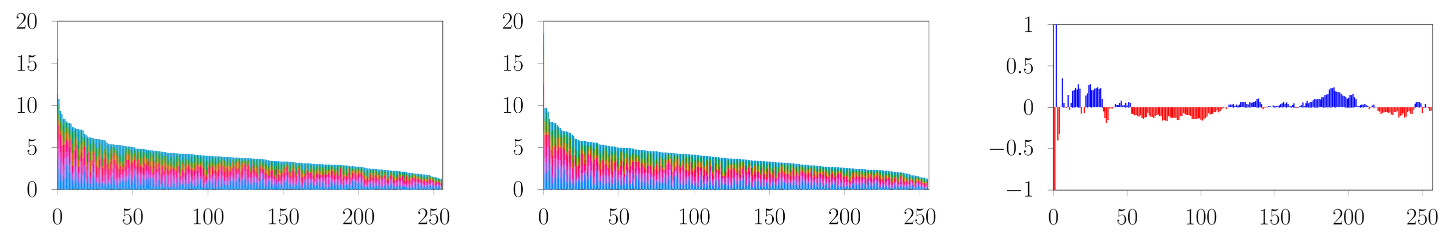
Merge: interactive semantic decoder

Visualizations

1. **SHAP values analysis:** DMAD maintains the SHAP values of the object query across two training stages, which **interprets the elimination of negative transfer**. From left to right: stage 1, stage 2, and the difference (stage 1 minus stage 2). In the difference diagram, red indicates a negative value and blue signifies a positive value.

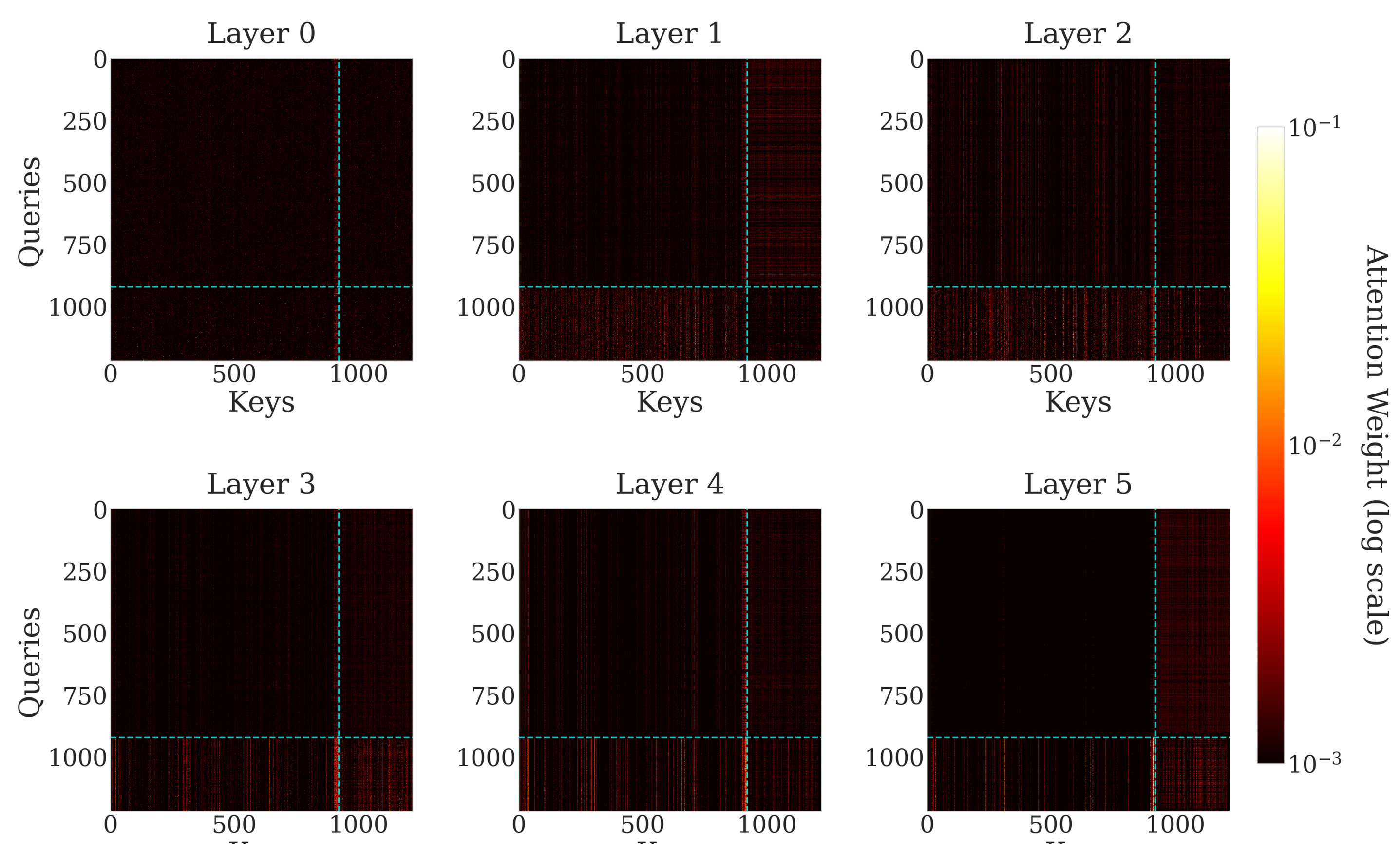


(a) SHAP values of UniAD.



(b) SHAP values of DMAD.

2. **Object & map self-attention heatmaps:** Cyan dashed lines divide a heatmap into four attention regions. **upper-left**: object to object; **upper-right** object to map; **lower-left**: map to object; **lower-right** map to map.



Results

Applying "Divide and Merge" structure to UniAD and SparseDrive, resulting in **DMAD** and **SparseDMAD**.

1. **Mitigating Negative Transfer:** Experiments on nuScenes benchmark show that our methods mitigate the negative transfer in training stage 2. Performance changes in stage 2 are shown in parentheses (red: decline, blue: improvement).

Method	NDS↑	mAP↑	AMOTA↑	AMOTP↓
UniAD - stage 1	0.497	0.382	0.374	1.31
UniAD - stage 2	0.491 (-1.2%)	0.377 (-1.3%)	0.354 (-5.3%)	1.34 (+2.3%)
DMAD - stage 1	0.504	0.395	0.394	1.32
DMAD - stage 2	0.506 (+0.4%)	0.396 (+0.3%)	0.393 (-0.3%)	1.30 (-1.5%)
SparseDrive - stage 1	0.531	0.419	<u>0.395</u>	<u>1.25</u>
SparseDrive - stage 2	0.523 (-1.5%)	0.417 (-0.5%)	0.376 (-4.8%)	1.26 (+0.8%)
SparseDMAD - stage 1	**0.536**	<u>0.424</u>	**0.396**	**1.23**
SparseDMAD - stage 2	<u>0.534</u> (-0.4%)	**0.427** (+0.7%)	<u>0.395</u> (-0.3%)	**1.23** (0%)

2. **Closed-loop planning:** Experiments on NeuroNCAP demonstrate that our advances in perception transform to planning safety.

Method	NeuroNCAP scores ↑				Collision rates (%) ↓			
	Stat.	Frontal	Side	Avg.	Stat.	Frontal	Side	Avg.
UniAD	3.50	1.17	1.67	2.11	32.4	77.6	71.2	60.4
DMAD	4.40	1.47	2.07	2.65	**14.8**	74.0	61.6	50.1
SparseDrive	4.42	2.96	2.30	3.23	22.4	62.8	60.4	48.5
SparseDMAD	**4.57**	**3.14**	**2.42**	**3.37**	18.4	**60.0**	**59.1**	**45.8**

Scan QR codes for paper & code

Paper **Code**