Recent Trends in E2E

RILS LAB

Robot Intelligence Learning & System

김채은

2026.08.07

01

GuideFlow: Constraint-Guided Flow Matching for Planning in End-to-EndAutonomous Driving

Long Nguyen, Micha Fauth, Bernhard Jaeger, Daniel Dauner, Maximilian Igl Andreas Geiger, Kashyap Chitta

CVPR / 2026

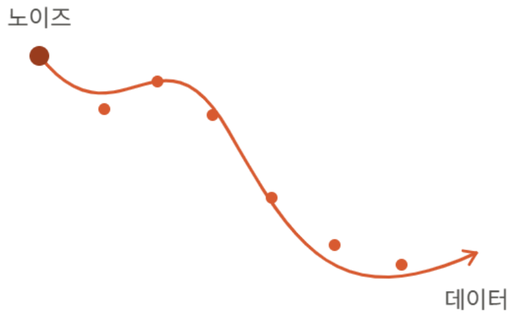
1. Problem Setting

Flow Matching이란?

노이즈에서 데이터로 가는 "속도장(velocity field)"을 단순 회귀로 학습하는 생성 모델. Diffusion처럼 확률적 노이즈 제거를 반복하는 게 아니라, ODE를 따라 흘려보냄

Diffusion 방식

노이즈를 조금씩 지워 나감



궤적이 휘어 있어 잘게 쪼개야 함

확률적 denoising

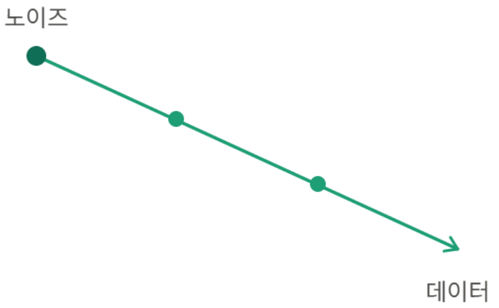
매 step 노이즈를 조금 제거

경로가 굽어 step을 많이 밟음

보통 50~1000 step

Flow matching 방식

노이즈에서 데이터로 곧장 흐름



dx/dt = v(x, t) 를 그대로 따라감

속도장 ODE 적분

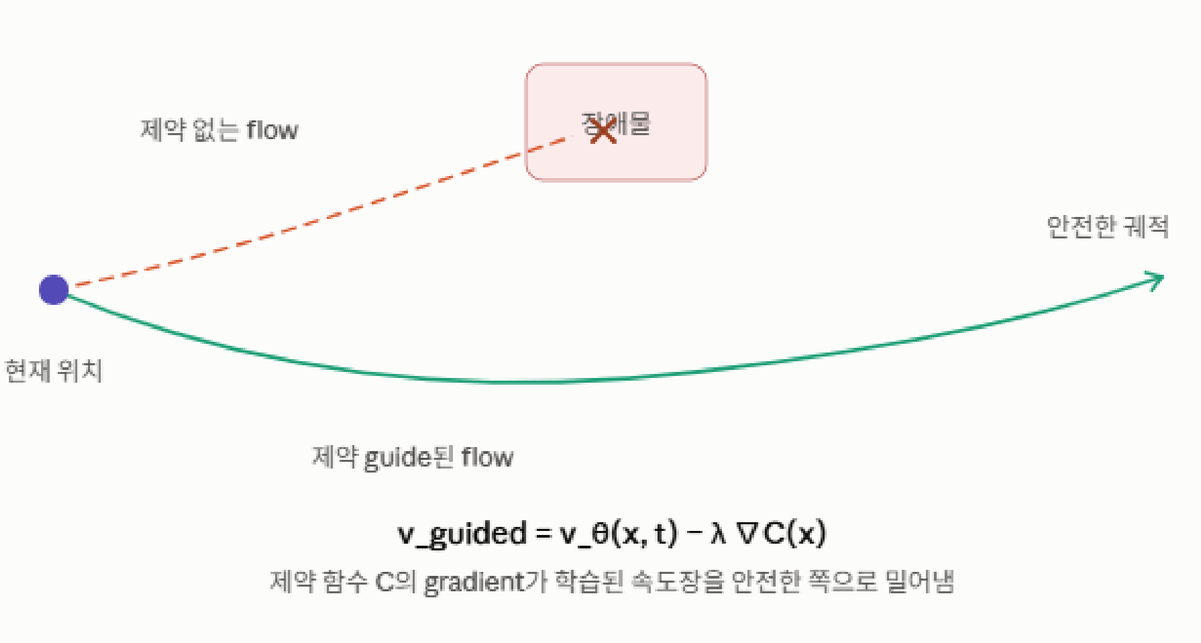
속도 v를 예측해 그쪽으로 이동

직선이라 큰 step도 오차 작음

2~10 step이면 충분

1. Problem Setting

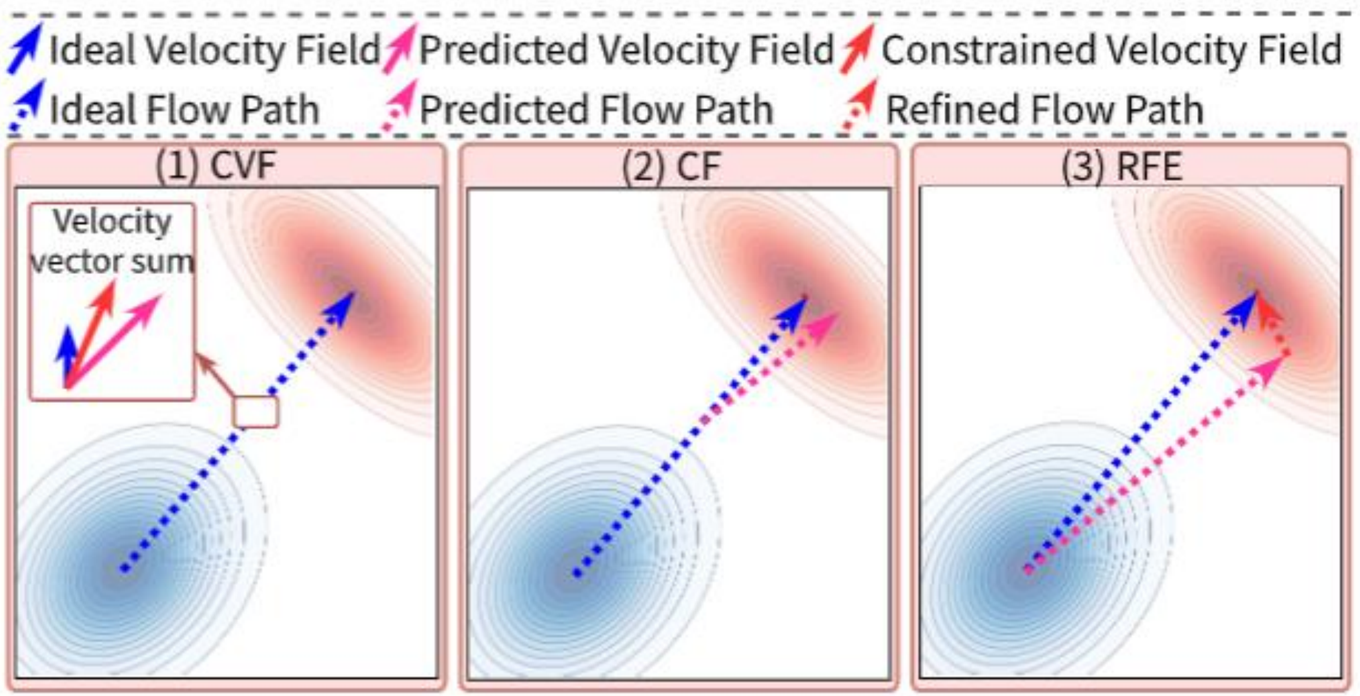
기존의 IL기반 E2E-AD는 multimodal trajectory mode collapse로 인해 다양한 궤적을 나타내지 못하는 경우가 많다.
그러나 생성형 E2E planner(generative E2E planner)는 생성과정에서 중요한 안전 및 물리적 제약조건을 통합하는데 어려움을 겪으며 결과물 개선을 위해 추가적인 최적화 단계가 필요함
본 논문에서는 제약조건이 guide된 흐름 매칭(Constrained Flow Matching)을 활용하는 새로운 Planning framework인 GuideFlow를 제안



2. Solution= CVF

BEV를 만들고 agent token과 map token을 생성한다.
랜덤한 경로를 X_0으로 만들어두고 두 쿼리와의 cross-attentio을 통해 상황에 대한 context를 주입하고 이를 MLP에 통과해 속도 필드 $v\theta(xt,t)$를 예측 함
이를 사전학습된 CVF(제약을 만족하는 속도 필드)와의 내적을 통해 보정된 속도 필드를 사용함
강력하지만 모델의 생성 dynamics에 가장 직접적으로 개입하는 방식이라 위험할 수 있음

$$v_t^* = v_t - 2\lambda \frac{v_t \cdot v_t^c}{\|v_t^c\|^2} v_t^c$$



2. Solution= Constrained Generation

BEV를 만들고 agent token과 map token을 생성한다.
랜덤한 경로를 X_0으로 만들어두고 두 쿼리와의 cross-attentio을 통해 상황에 대한 context를 주입하고 이를 MLP에 통과해 속도 필드 $v\theta(xt,t)$를 예측 함

- **CVF**
  - 현재 이동 방향이 잘못되는 것을 즉시 완화
- **CF**
  - generation 중 누적되는 trajectory drift를 직접 보정
- **RFE**
  - 최종 결과가 constraint를 만족하는 low-energy mode에 도달하도록 최적화

CVF 공식

$$v_t^* = v_t - 2\lambda \frac{v_t \cdot v_t^c}{\|v_t^c\|^2} v_t^c$$

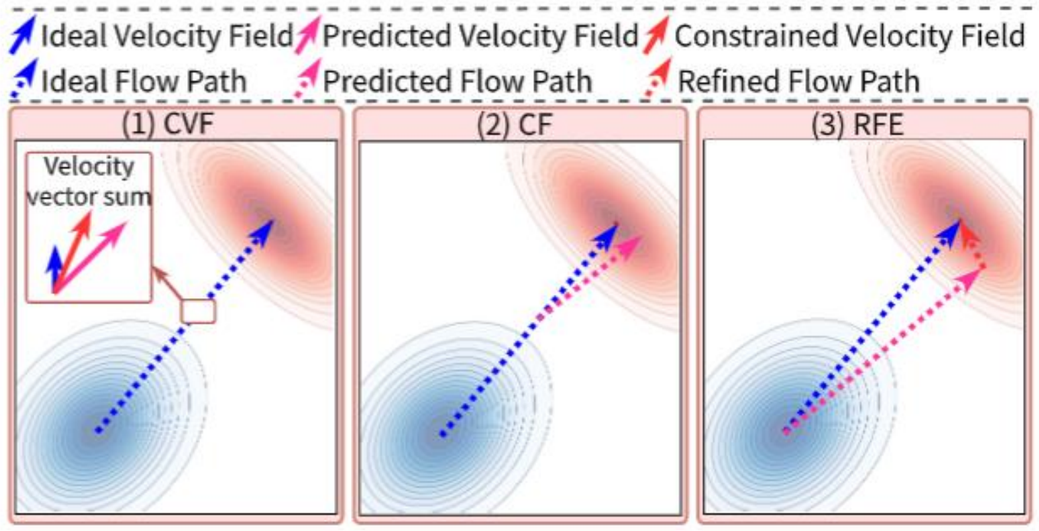
CF 공식

$$\phi'_t = \{x^{(0)}, x^{(1)}, \ldots, x^{(K)}\}$$

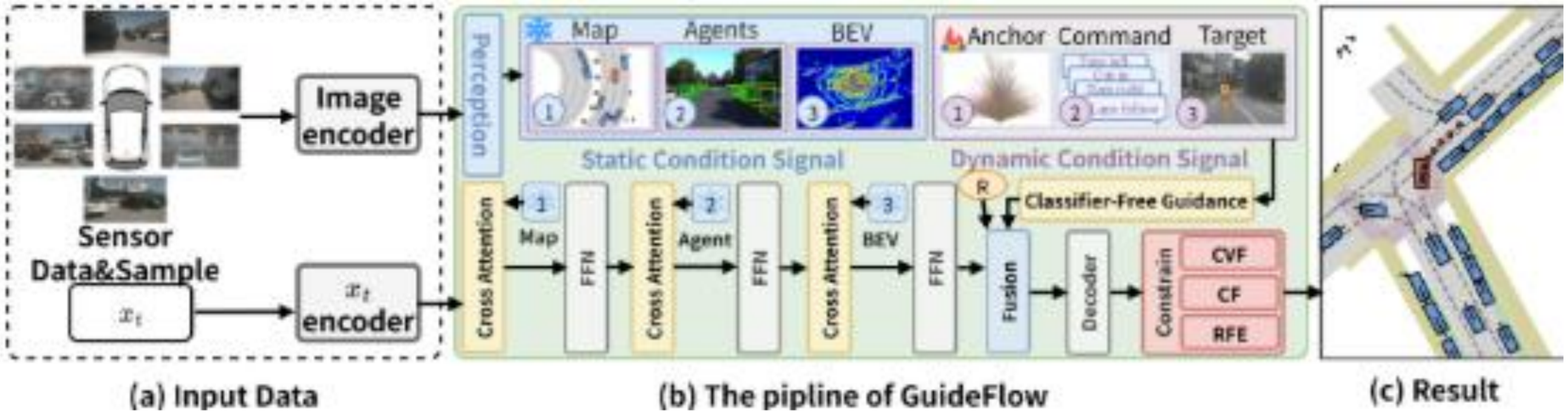
$$x^{(k_c)} = x_1^c$$

REF 공식

$$E_\theta(x_t) = \left\| \mathrm{J}\left(f_{t>1}(x_t)\right) - \mathrm{J}(x_t) \right\|^2$$



3. Contribution



1. Flow Matching이 여러 초기 noise에서 다양한 trajectory 후보를 생성
2. CVF가 각 step의 방향을 안전한 방향으로 조정
3. CF가 필요할 경우 중간 flow state를 안전한 anchor로 교체
4. RFE가 최종 단계에서 trajectory를 낮은 energy의 feasible region으로 refinement

3. Contribution

Table 5. Ablation studies of different modules in GuideFlow over NavSim [7] HavHard Split, Bench2Drive [17], NuScenes [2] and ADV-NuScenes [43]. "EP" stands for Ego Progress subscore. "CVF" denotes "Constraining the Velocity Field" module, "CF" denotes "Constraining the Flow State", "RFE" denotes "Refining the Flow by EBM" and "RAS" denotes "Reward as Style Condition".

Modules				NavSim Hard (No Scorer)				Bench2Drive		NuScenes	ADV-NuScenes
CVF	CF	RFE	RAS	EP	Stage1 EPDMS	Stage2 EPDMS	EPDMS	Drive Score	Success Rate	C.R (%)	C.R (%)
				84.1	56.7	40.0	23.1	73.86	50.00	0.08	1.02
✓				80.9	56.9	41.3	24.5	74.22	50.00	0.08	0.94
	✓			81.7	57.1	44.7	25.1	74.67	50.45	0.07	0.77
		✓		81.1	53.3	45.3	25.5	74.90	50.45	0.07	0.83
	✓	✓		79.6	54.9	**47.9**	**27.1**	**75.21**	**51.36**	**0.07**	**0.73**
	✓	✓	✓	82.3	**60.1**	43.7	26.3	74.46	50.45	0.08	0.80

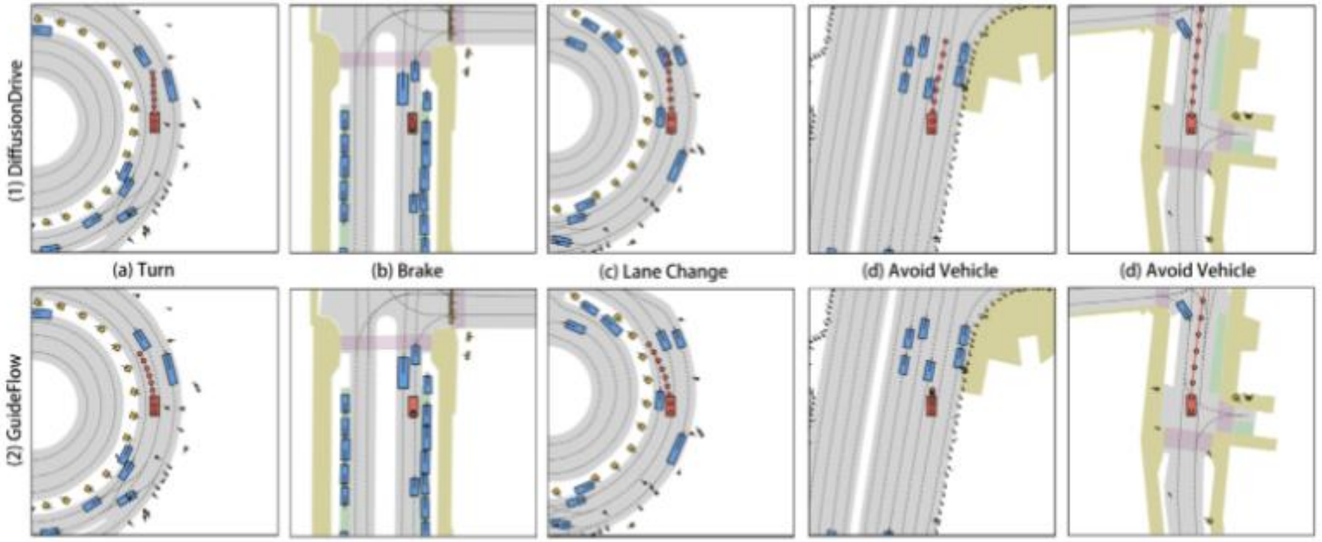


Figure 4. **Visual comparison between DiffusionDrive [27] and our GuideFlow across multiple driving scenarios.** GuideFlow generates trajectories that exhibit improved adherence to lane follow, smoother maneuver transitions, and stronger compliance with safety constraints such as collision avoidance and road boundary preservation.

Mode collapse를 줄이는 multimodal planner 제안

Constraint를 generation 과정에 직접 적용

CF

- 생성 trajectory를 직접 feasible한 상태로 끌어옵니다.
- 명시적인 constraint satisfaction에 강합니다.

RFE

- energy landscape를 통해 모델이 안전한 영역을 스스로 탐색하도록 합니다.
- constraint rule의 일반화와 OOD 대응에 강합니다.

02

DriveMoE: Mixture-of-Experts for Vision-Language-Action Model inEnd-to-End Autonomous Driving

Zhenjie Yang, Yilin Chai, Xiaosong Jia2,3∗Qifeng Li1,4 Yuqian Shao1,4 Xuekai Zhu1
Haisheng Su1 Junchi Yan1†

CVPR / 2026

1. Problem Setting

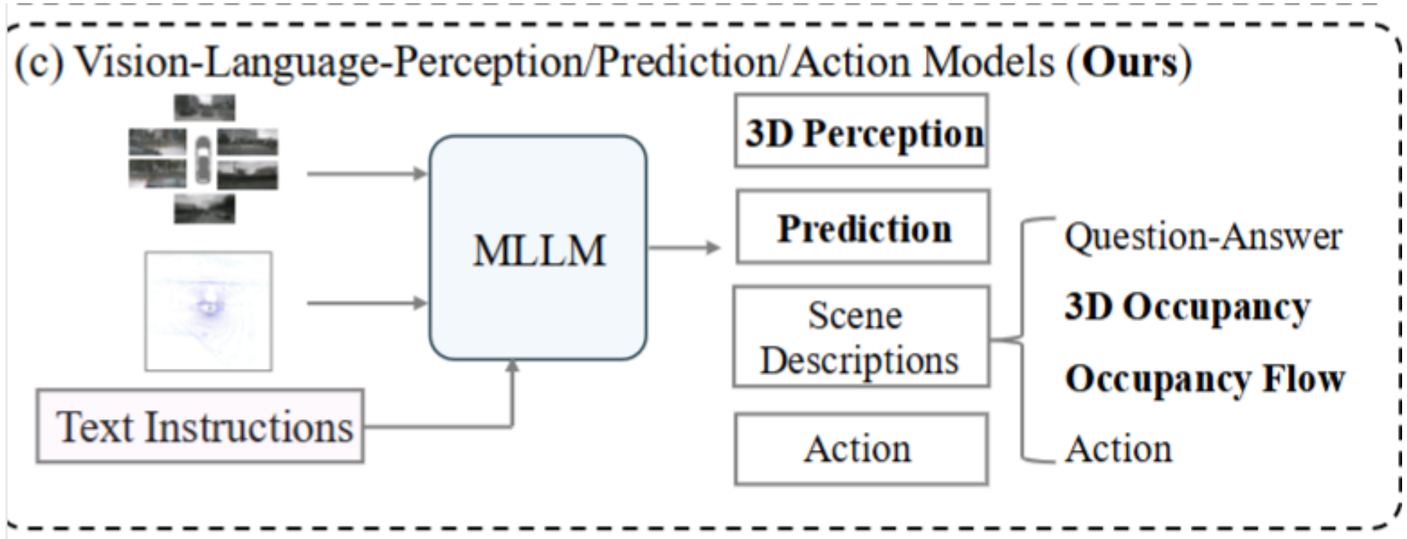
자율주행 차량은 여러 개 카메라의 temporal video를 동시에 사용해야하는데 기존 VLA모델은 여러카메라의 모든 이미지를 vision tower에 입력하고 생성된 모든 visual token을 언어모델에 전달하는 방식

➔ 카메라 수와 temporal frame이 비례해 transformer의 입력이 길어져 계산량과 메모리 사용량이 증가함

모든 시점에서 모든 카메라가 동일하게 중요할까????

2. Solution

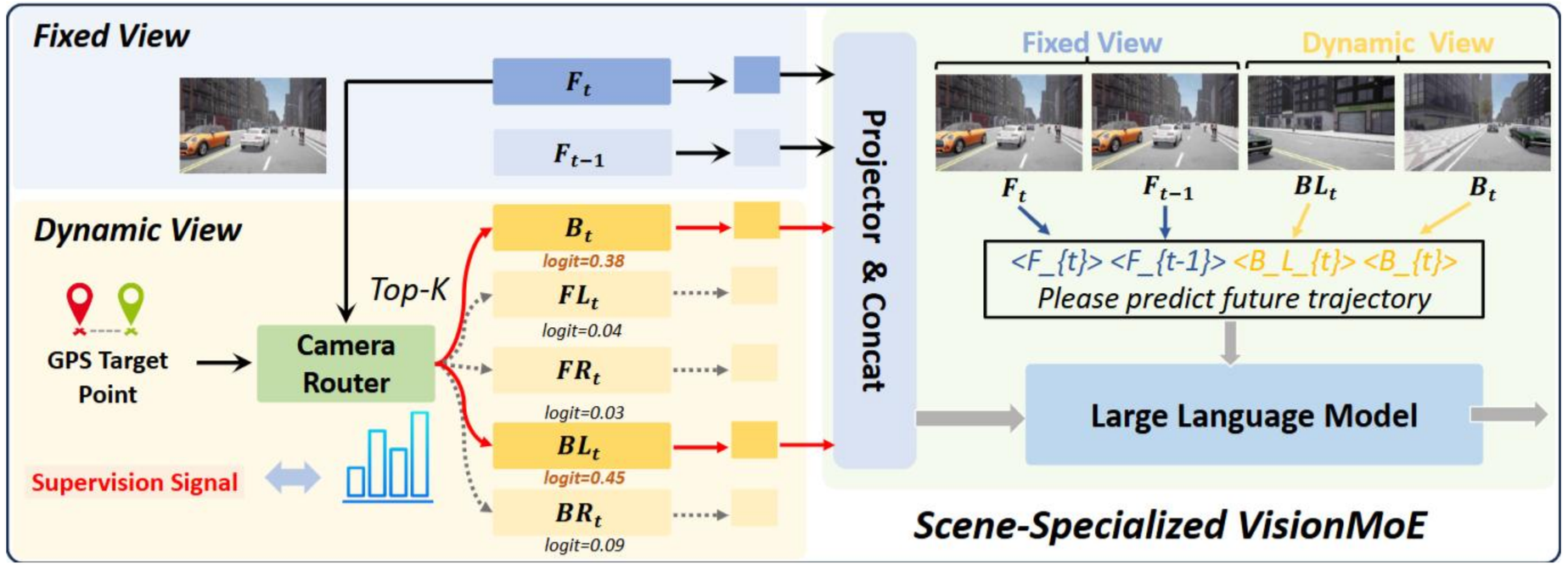
Drive-\(\pi_0\)는 Embodied AI용 VLA 모델인 \(\pi_0\)를 자율주행에 맞게 확장한 baseline이다.



INPUT
두 개의 연속된 전방 카메라 이미지
“Please predict future trajectory”와 같은 고정 text prompt
차량 상태 정보(속도, yaw, 과거 궤적)
이후 PaliGemma VLM이 시각 및 언어 정보를 인코딩하고,
flow-matching 기반 action module이 미래 궤적을 생성합니다.

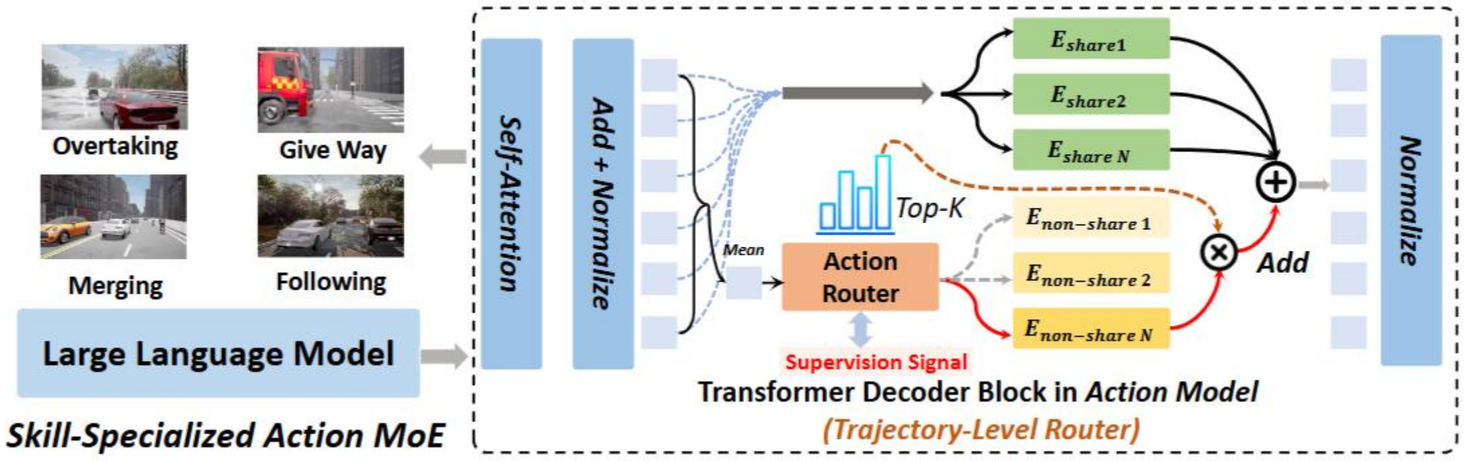
2. Solution:Vision MoE

Vision Router이 현재 장면과 주행 목표에 따라 중요한 카메라 view를 동적으로 선택합니다.
연속된 2개의 정면 카메라 프레임
1개의 동적으로 선택되는 현재 카메라 프레임
총 3개의 프레임만을 본다.

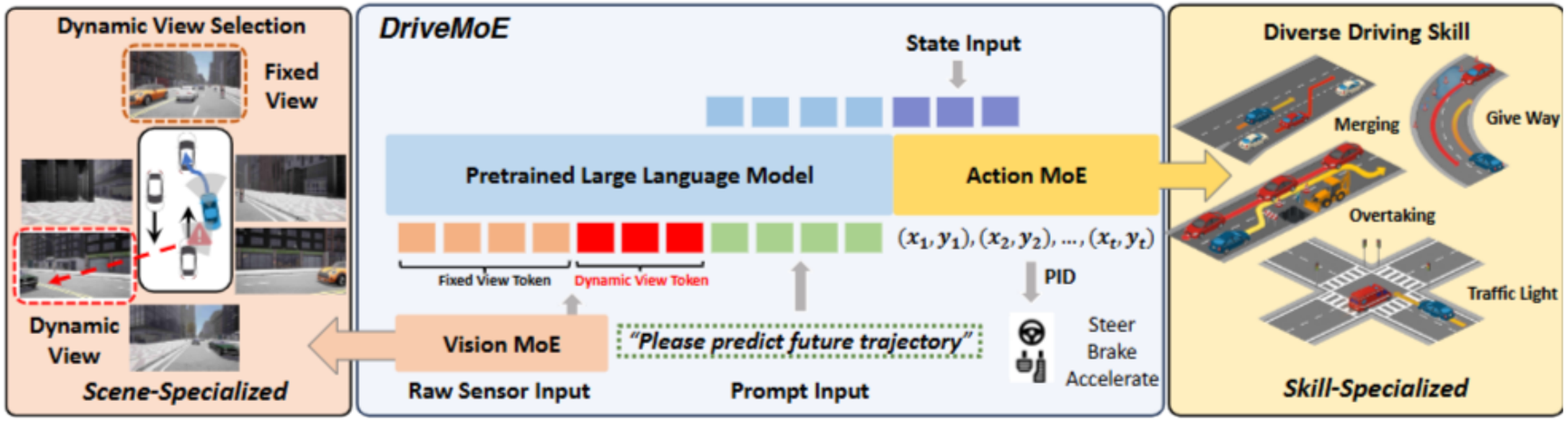


2. Solution: Action MoE

Action MoE는 하나의 FFN 또는 정책 네트워크가 모든 운전 행동을 처리하는 대신, 여러 개의 expert FFN을 두고 현재 trajectory 또는 driving skill에 적합한 expert를 선택한다. 가장 일반적인 주행을 하는 shared exper를 모든 궤적 생성에 공통으로 사용하고 6개의 non-share에서 top 3개를 선택해 가중합한다.



2. Solution: PipeLine



3. Contribution

Method	Venue	Ability (%) ↑					
		Merging	Overtaking	Emergency Brake	Give Way	Traffic Sign	Mean
TCP-traj* [54]	NeurIPS 2022	8.89	24.29	51.67	40.00	46.28	34.22
AD-MLP [61]	Arxiv 2023	0.00	0.00	0.00	0.00	4.35	0.87
UniAD-Base [16]	CVPR 2023	14.10	17.78	21.67	10.00	14.21	15.55
ThinkTwice* [23]	CVPR 2023	27.38	18.42	35.82	**50.00**	54.23	37.17
VAD [31]	ICCV 2023	8.11	24.44	18.64	20.00	19.15	18.07
DriveAdapter* [22]	ICCV 2023	28.82	26.38	48.76	**50.00**	56.43	42.08
DriveTrans [26]	ICLR 2025	17.57	35.00	48.36	40.00	52.10	38.60
DiffAD [52]	Arxiv 2025	30.00	35.55	46.66	40.00	46.32	38.79
Drive-π_0 (Ours)	Ours	26.25	26.67	45.00	30.00	38.95	33.37
DriveMoE (Token-Level)		**28.75**	**31.11**	**51.67**	40.00	**52.63**	**40.83**
DriveMoE (Traj-Level)		**34.67**	**40.00**	**65.45**	40.00	**59.44**	**47.91**

Vision MoE를 이용한 동적 카메라 선택

Action MoE를 이용한 운전 skill 전문화

이러한 방식을 통해 최근 모델중 높은 성적 달성

03

AutoDrive-P³: Chain of Perception–Prediction–Planning via RF

Yuqi Ye, Zijian Zhang, Junhong Lin, Shangkun Sun, Changhao Peng, Wei Gao

ICLR / 2026

1. Problem Setting

기존의 작은 E2E모델은 long-tail상황에 약해 world knowledge가 있는 VLM을 모델에 붙힘

그러나 저자는 이 방식에 3가지 문제가 있다고 봄

1. CoT없이 궤적을 바로 뱉는다.
2. Perception,prediction,plannin을 따로 답함
3. GRPO를 plannin에만 건다.

3단계 모두에서 GRPO를 걸어 연관성을 줄 수 없을까?

2. Solution: reward funtion

① **Perception reward** — 객체 탐지 품질

$$R_{perc} = \mathrm{IoU}_{avg} \cdot (0.5P + 0.5R)$$

위치 정확도(IoU)와 검출 균형의 곱. 예외적으로 **GT도 없고 예측도 없으면 1.0**을 줍니다.
한쪽만 비면 0.0
"아무것도 없는 장면에서 아무것도 없다고 말하는 것"도 정답으로 인정

② **Prediction reward**

$$R_{pred} = \underbrace{\left(\frac{\sum_{(i,j)\in\mathcal{M}} \mathrm{IoU}_{ij} \cdot \mathbb{I}(s_i = s_j)}{\sum_{(i,j)\in\mathcal{M}} \mathrm{IoU}_{ij}} \right)}_{\text{IoU로 가중된 행동 라벨 정확도}} \times \underbrace{\left(\mathrm{IoU}_{avg} \cdot (0.5P + 0.5R) \right)}_{= R_{perc}\text{와 동일한 항}}$$

λperc:λpred:λplan=1:2:2:5

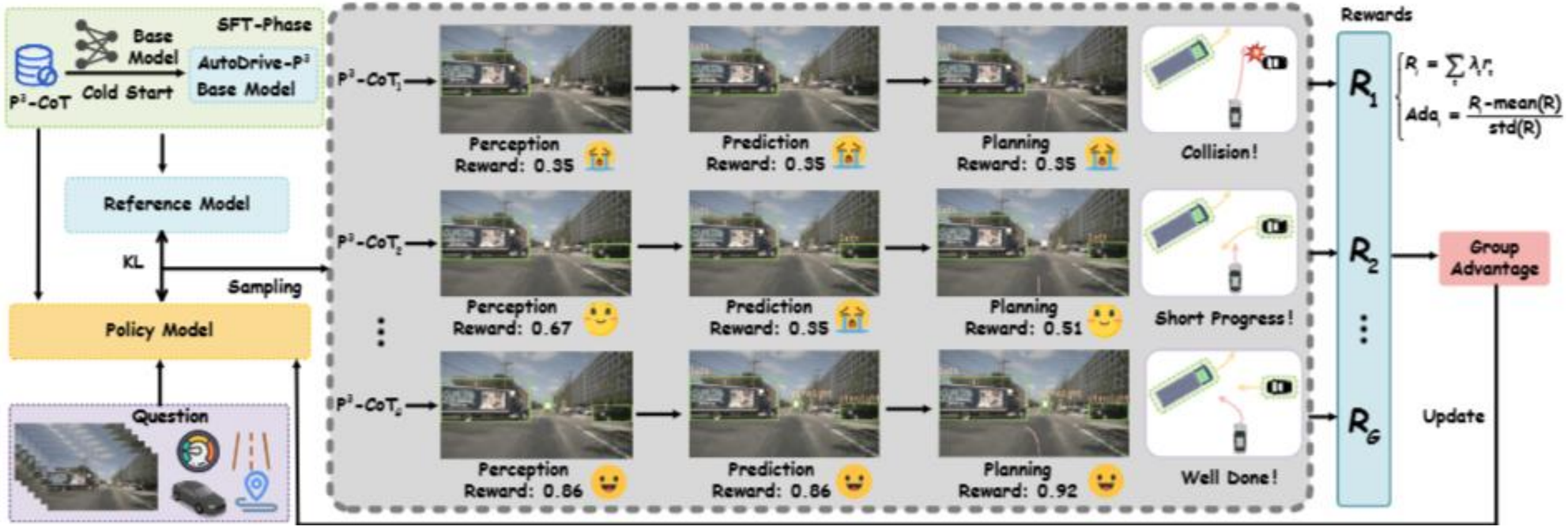
③ **Planning reward** — 궤적 품질

$$R_{plan} = \frac{2}{1 + e^{\mathrm{clip}(L2, 0\, L2_{max})}}$$

L2 거리를 시그모이드 형태로 뒤집은 것. L2=0이면 1.0이고 멀어질수록 0으로 수렴.
clip은 아주 틀린 궤적이 보상을 완전히 죽이는 걸 막는다.

2. Solution: pipeline

Supervised Fine-Tuning(SFT): 모델을 자율주행 도메인에 정렬하기 위해, 단계별 추론 y_thinking과 답변y_answer을 생성하도록 학습



3. contribution

Table 2: **Performance comparison on NAVSIMv1 benchmark.**

Method	Image	Lidar	NC↑	DAC↑	EP↑	TTC↑	Comf↑	PDMS↑
Human	✗	✗	100.0	100.0	87.5	100.0	99.9	94.8
Constant Velocity	✗	✗	69.9	58.8	49.3	49.3	100.0	21.6
Ego Status MLP	✗	✗	93.0	77.3	62.8	83.6	100.0	65.6
VADv2 (Weng et al., 2024)	✔	✗	97.9	91.7	77.6	92.9	100.0	83.0
UniAD (Hu et al., 2023)	✔	✗	97.8	91.9	78.8	92.9	100.0	83.4
LTF (Prakash et al., 2021)	✔	✗	97.4	92.8	79.0	92.4	100.0	83.8
TransFuser (Prakash et al., 2021)	✔	✔	97.7	92.8	79.2	92.8	100.0	84.0
PARA-Drive (Weng et al., 2024)	✔	✗	97.9	92.4	79.3	93.0	99.8	84.0
LAW (Li et al., 2024a)	✔	✔	96.4	95.4	81.7	88.7	99.9	84.6
DRAMA (Yuan et al., 2024)	✔	✔	98.0	93.1	80.1	94.8	100.0	85.5
Hydra-MDP (Li et al., 2024b)	✔	✔	98.3	96.0	78.7	94.6	100.0	86.5
DiffusionDrive (Liao et al., 2025)	✔	✔	98.2	96.2	82.2	94.7	100.0	88.1
WoTE (Li et al., 2025)	✔	✔	98.5	96.8	81.9	94.9	99.9	88.3
AutoDrive-P^3 (Ours-Detailed)	✔	✗	**99.1**	97.4	**84.8**	96.5	**100.0**	**90.6**
AutoDrive-P^3 (Ours-Fast)	✔	✗	98.9	**97.7**	83.7	**96.6**	99.9	90.2

라이다 없이 카메라만 가지고 우수한 성적
강화학습을 인지,판단,제어단에 다 적용해 각각의 관계도 반영한 것

04

LEAD: Minimizing Learner–Expert Asymmetry in End-to-End Driving

Long Nguyen, Micha Fauth, Bernhard Jaeger, Daniel Dauner, Maximilian Igl, Andreas Geiger, Kashyap Chitta

CVPR / 2026

1. Problem Setting

Simulator에서 GT를 ptiviled expert를 통해 만드는데 이는 내부 상태를 다 읽어버리는 rule base 프로그램이다.

- 모든 차량의 정확한 위치, 속도, 신호등상태, 벽 뒤에 숨은 차 등을 전부 API로 꺼내 쓴다.

Expert가 벽 뒤에 가려지거나 보이지 않는 것에 반응해 브레이크를 밟으면 student는 화면에 아무것도 없는데 감속을 해야한다.(non-casual demonstration)

- Expert는 다른 차의 속도,가속도를 오차 0으로 안다.
  그러니 아슬아슬한 주행이 가능하지만 student는 추정치로 판단하니 성능을 내기가 어렵다.
- Expert는 전체 경로를 다 알지만 student는 보통 target point하나만 안다.

정보다 턱없이 부족헤 정책이 이 target point에 과하게 매달리는 target point bias가 생김

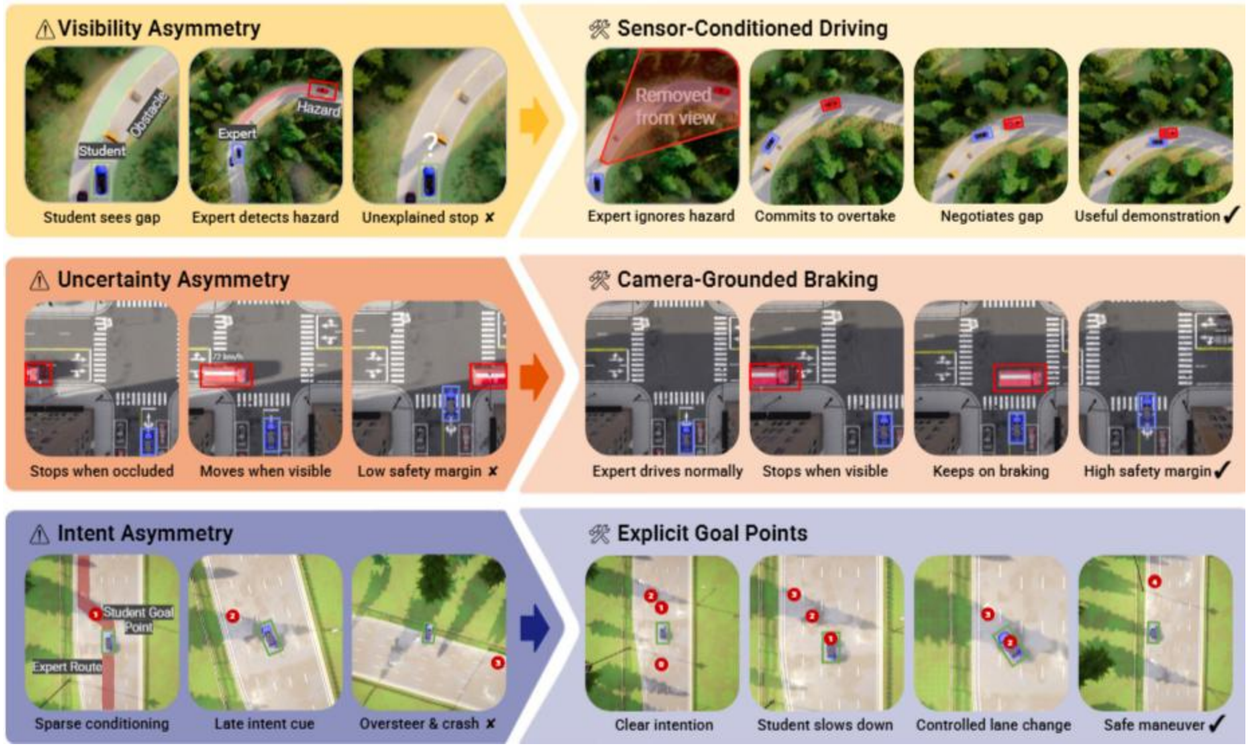
Expert를 약하게 만들어 학생이 따라할 수 있는 정답으로 만들자

2. Solution: PipeLine

Expert plannig의 입력을 student 카메라 frustum안에 들어오는 것으로 제한한다.
딱 student가 볼 수 있는 만큼 본다. 가려진 차를 못 본다,
속도제한도 API가 아닌 카메라로 추론 가능한 단서에서 목표속도를 정하도록 설정한다.

불확실설 비대칭 처방
1. 정밀한 속도, 가속도 추정에 의존 대신 가까이 있거나 실제 보이는 위험 대상에 브레이크를 건다.
2. 야간이나 악천후처럼 시야가 나쁜 조건에는 주행속도 자체를 낮춘다.
3. 비보호 좌회전 같은 상황에서 마주오는 차의 bounding box를 실제보다 크게 부풀린다.

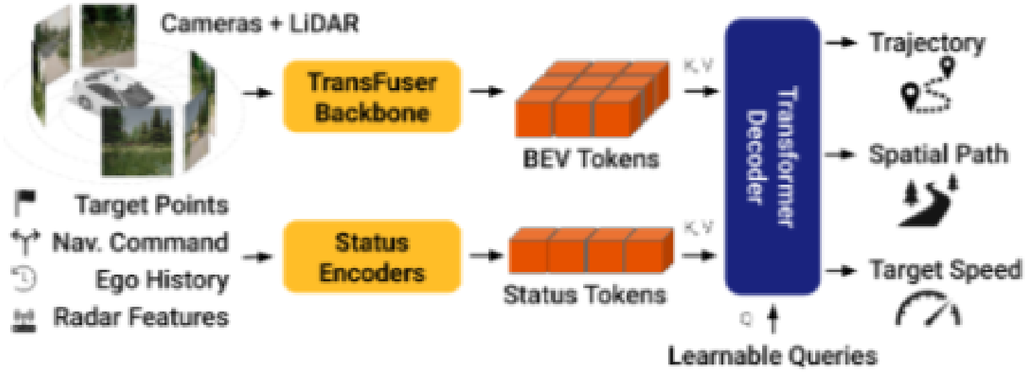
Method	Backbone	Input Modalities			Bench2Drive	
		Camera FOV	LiDAR	Radar	Driving Score ↑	Success Rate ↑
HiP-AD [52]	ResNet-50	360°	✗	✗	86.8	69.1
SimLingo [45]	InternViT-300M	110°	✗	✗	85.1	67.2
TFv5 [64]	RegNetY-032	110°	✓	✗	83.5 + 0.3	67.3 + 1.0
TFv6 (Ours)	ResNet-34	360°	✗	✗	91.6 + 0.7	79.5 + 2.0
		360°	✓	✗	94.7 + 0.6	85.6 + 0.0
		360°	✗	✓	94.2 + 0.7	85.3 + 0.9
		360°	✓	✓	95.0 + 0.7	84.3 + 2.1
		140°	✓	✓	94.7 + 0.7	82.1 + 3.6
	RegNetY-032	140°	✓	✓	**95.2** + 0.3	**86.8** + 0.7
PDM-Lite [51]	-		-		*97.0*	*92.3*
LEAD (Ours)	-		-		*96.8*	*96.6*



2. Solution: student model

① GRU 제거
기존 모델은 planning query를 target point 정보로 다듬을 때 **GRU**를 사용.
문제는 이게 파이프라인 **끝단**에 붙어 있어 목표 지점 정보가 모든 인지 처리가 끝난 뒤에야 주입됨 (late goal conditioning).
논문은 이 GRU를 **얕은 병목(shallow bottleneck)** 으로 지목
좁은 통로를 통과하면서 target point의 영향력만 증폭되고, 그 결과 정책이 경직됨
② target point를 명시적 토큰으로
GRU 자리를 대신해 삽입
target point를 [-1, 1]로 정규화한 **토큰 하나**로 만들어서, planning decoder에서 **BEV 토큰들과 나란히** 놓음
attention을 통해 인지 feature와 주행 의도가 훨씬 이르게, 그리고 풍부하게 섞임
핵심 전환= 목표 지점이 **마지막에 결과를 비트는 신호**에서 **처음부터 장면과 함께 해석되는 재료**로 변함
③ 경로 조건을 촘촘하게
점 하나 대신 **이전 / 현재 / 다음 세 점**으로 경로를 표현. 그리고 현재 target point를 다음 점으로 넘기는 거리 임계값도 줄임.
그러면 목표점이 더 자주 갱신되면서 방향 정보가 끊기지 않음.

백본은 RegNetY-032, 레이더도 지원하는데 레이더는 객체 수준 feature로 뽑아서 planning decoder에 추가 key/value 토큰으로 그냥 밀어 넣습니다.



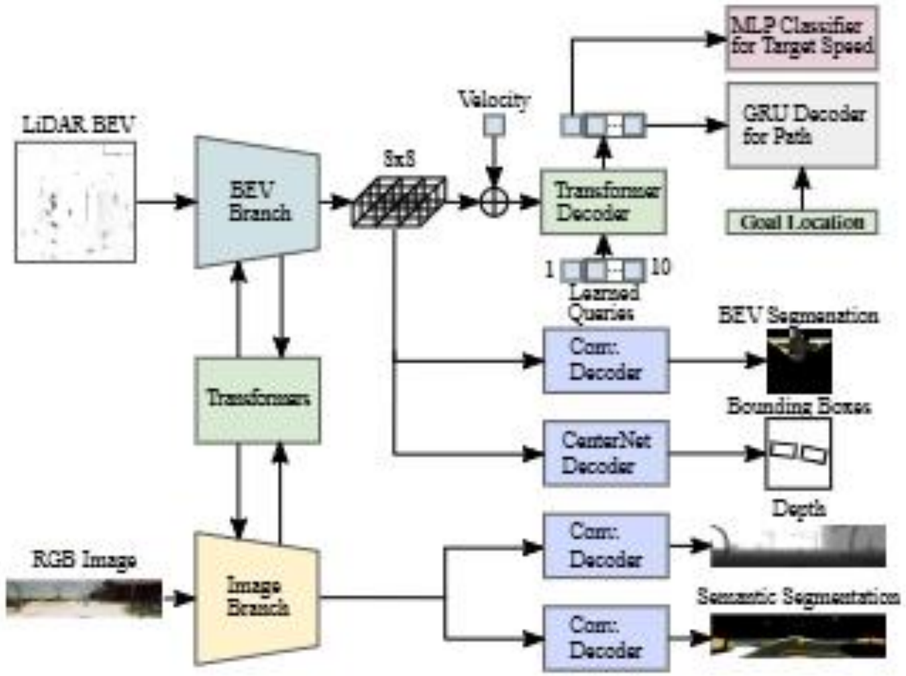


Figure 7: **TransFuser++ architecture.**

3. contribution

Method	RC ↑	IS ↑	DS ↑	I ↑	NDS ↑
'Town13 Val' – Town13 withheld during training					
TFv5 [64]	**50.20**	0.10	1.08	0.04	2.12
TFv6 (Ours)	39.70	**0.30**	**3.52**	**0.22**	**4.04**
PDM-Lite [51]	*83.40*	*0.41*	*36.30*	*0.63*	*58.50*

Learner–Expert Asymmetry를 체계적으로 분석하고 완화함

TFv6라는 개선된 TransFuser pipeline을 제시함

Thank you

Question & Disscusion

RILS LAB