

CAR Planner : 제약 어텐션 기반 강건 모방학습 플래너

김지윤 ¹⁾·최경환*²⁾

광주과학기술원 ¹⁾·한국과학기술원 ²⁾

CAR Planner: Constrained-Attention-Based Robust Imitation Learning for Autonomous Driving

Jiyeon Kim¹⁾·Kyunghwan Choi*²⁾

¹⁾Artificial Intelligence Graduate School, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea

*²⁾Cho Chun Shik Graduate School of Mobility, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea

Abstract : Imitation Learning (IL) is widely used for autonomous driving but remains vulnerable to shortcut learning, where policies exploit spurious correlations rather than causal driving cues, undermining generalization and robustness under distribution shift and OOD conditions. We present CAR Planner, which augments the attention mechanism with a constrained formulation: an inequality constraint limits the mean deviation of ego-state cross-attention weights from uniform, enforced via the Augmented Lagrangian Method (ALM). This regularizer curbs over-reliance on a few channels and encourages balanced state representations, adding negligible training overhead and no inference cost. On nuPlan, CAR Planner shows markedly smaller performance drops when ego-state channels are reduced, providing strong evidence of robustness against shortcut reliance. Furthermore, it consistently outperforms both an unconstrained baseline(without constraint) and state-of-the-art dropout-based methods in challenging scenarios, while also producing smoother and more stable driving. These results indicate that CAR Planner enables more robust IL for autonomous driving.

Key words : Attention(어텐션), Autonomous driving(자율주행), Constraint optimization(제약 최적화), Imitation learning(모방학습), Augmented lagrangian method(증강 라그랑지안 방법)

1. 서론

현실 세계의 다양한 상황에서도 강건성을 유지하는 인간 수준의 모션 플래닝은 자율주행의 핵심 과제이다. 데이터 기반 접근 중 모방학습(Imitation Learning, IL)은 전문가 시연을 모사하는 방식으로 널리 활용되고 있다.¹⁻³⁾ 그러나 IL은 두 가지 주요 한계를 가진다. 첫째, 오차 누적(compounding error)은 작은 편차가 시간이

지남에 따라 확대되며 결국은 주행 실패로 이어지는 문제로, 노이즈 주입이나 도메인 랜덤화 같은 데이터 증강으로 완화할 수 있다.^{4,5)} 둘째, 쇼트컷 학습(shortcut learning)으로, 정책이 인과적 주행 행태 대신 학습 데이터의 우연한 상관관계에 집중하는 현상이다.^{6,7)} 이는 분포 변화나 OOD(out-of-distribution) 상황에서 강건성을 크게 저하시킨다.

nuPlan 벤치마크⁸⁾에서는 이 문제가 특히 ego 상태(ego-state) 표현에서 두드러진다. 플래너가

* Kyunghwan Choi, E-mail: kh.choi@kaist.ac.kr

속도나 조향각 등 일부 특성에 과도하게 집중하면, 개방형(open-loop)에서는 높은 점수를 얻지만 폐루프(closed-loop)에서는 성능이 급격히 떨어진다. 이는 어텐션 붕괴(attention collapse)로 나타나며, 특정 채널에 어텐션이 과도하게 쏠리는 현상이다. 기존 해결책으로는 패널티 정규화나 SDE(State Dropout Encoder)⁹⁾가 있으나, 전자는 태스크 손실과 간섭하며 튜닝이 어렵고, 후자는 입력 일부를 버리므로 일관된 성능을 보장하기 어렵다.

이에 본 연구는 CAR Planner를 제안한다. 제안 방법은 ego 상태 교차어텐션 가중치의 균등분포 대비 편차에 부등식 제약을 두어 특징 활용의 균형을 보장한다. 제약은 학습 단계에서 증강 라그랑지안 방법(ALM)^{10,11)}으로 집행되며, 입력 정보를 유지하면서 어텐션 붕괴를 방지한다. nuPlan 실험을 통해 CAR Planner의 Shortcut 완화 효과와 강건성 향상을 검증하며, OLS와 RCLS 지표에서 기존 방법 대비 일관된 개선을 보였다.

2. 관련 연구

초기 자율주행의 규칙 기반 플래너는 IDM¹²⁾, PDM-Closed¹³⁾ 등과 같이 규칙과 휴리스틱에 의존하여 해석 가능성과 정상 상황에서의 안전성을 확보했으나, 분포 변화나 새로운 시나리오에는 취약했다. 이후 UrbanDriver¹⁴⁾, GC-PGP¹⁵⁾, PDM-Open¹³⁾과 같은 학습 기반 플래너가 등장해 데이터로 정책을 직접 학습하며 성능을 높였지만, 학습 데이터의 우연한 상관관계에 과적합되는 쇼트컷 학습 문제로 인해 강건성이 저하되는 한계가 있었다. 이를 해결하기 위해 트랜스포머 기반 IL 플래너에서는 어텐션을 활용해 이고 상태, 주변 에이전트, 맵 정보를 통합하며, nuplan의 SOTA 방법인 planTF의 SDE처럼 드롭아웃으로 의존도를 낮추는 방법이 시도되었다.⁹⁾ 그러나 입력의 일부를 버리는 특성 때문에 분포 이동이 심한 상황에서는 성능 일관성에 한계가 있었다. 본 연구는 입력을 온전히 보존하면서 제약 최적화 방법 증강 라그랑지안 방법(ALM)을 통해 어텐션 분산을 직접 제어하여 쇼트컷 의존을 줄이고자 한다.

한편, 규칙과 학습을 결합한 하이브리드 접근이나 강화학습·페루프 학습 전략도 강건성을 높일 수 있으나 시스템 복잡도를 증가시키는 단점이 있다. 이에 본 연구는 순수 IL 틀에서 원리 기반 어텐션 정규화만으로 강건성을 향상시키는 방법을 제안한다.

3. 제안 방법

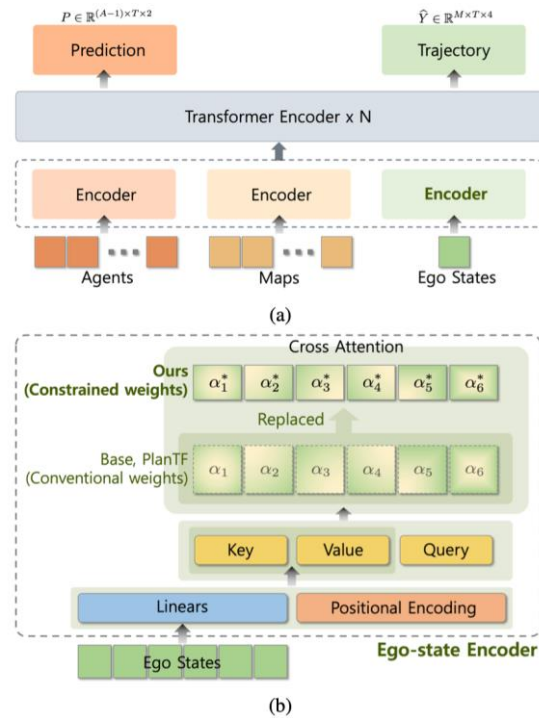


Fig.1 Base(바닐라), PlanTF(SDE 적용), Ours: CAR Planner(제약된 어텐션 가중치)의 플래너 아키텍처: (a) 전체 아키텍처 개요, (b) ego 상태 인코더.

CAR Planner는 Figure 1과 같이 트랜스포머 인코더-디코더 구조 위에서 ego 상태(6 채널) 교차어텐션의 분산을 직접 제약해 쇼트컷 의존을 완화한다. 입력은 (i) 주변 에이전트의 과거 상태, (ii) 현재 ego 상태 (x, y, yaw, v, a, s), (iii) 차선 폴리곤 기반 로컬 맵으로 구성한다. 에이전트와 맵 폴리곤은 전용 인코더로 먼저 임베딩한다. Ego를 포함한 모든 토큰에 시점·위치·헤딩 정보를 투영해 공용 트랜스포머 인코더로 통합한다. 인코딩된 ego 토큰으로부터, 궤적

디코더가 길이 길이 T 의 후보 궤적 M 개와 각 후보의 로짓을 생성하고, 최종 선택된 궤적을 출력한다. 에이전트의 미래 위치는 앞은 MLP로 $T \times 2$ 좌표 시퀀스를 예측한다. 이때 $T = 80, M = 6$ 이다.

먼저 imitation loss는 선택된 ego 궤적 가설 \hat{t} 에 대해 회귀 항과 모드 분류 항을 더해 아래와 같다.

$$\mathcal{L}_{reg} = L1_{smooth}(\hat{t}, Y) \quad (1)$$

여기서 $Y \in \mathbb{R}^{T \times 4}$ 는 ego 궤적의 GT로, 각 시점의 $(x, y, \cos\psi, \sin\psi)$ 를 포함한다.

$$\mathcal{L}_{cls} = CE(\pi, y^{mode}) \quad (2)$$

$$\mathcal{L}_i = \mathcal{L}_{reg} + \mathcal{L}_{cls} \quad (3)$$

에이전트 예측 손실은 다음과 같다.

$$\mathcal{L}_p = L1_{smooth}(P, P^{gt}) \quad (4)$$

여기서 $P, P^{gt} \in \mathbb{R}^{A-1 \times T \times 2}$ 는 에이전트 별 미래 위치의 예측/GT 시퀀스이며, 유효 시점과 에이전트에 대해 $L1_{smooth}$ 를 평균한다.

최종 태스크 손실은 아래와 같다.

$$\mathcal{L}_{task} = \mathcal{L}_i + \mathcal{L}_p \quad (5)$$

정리하면,

$$\mathcal{L}_{task}(\theta) = L1_{smooth}(\hat{Y}_{m^*}, Y) + CE(\pi, m^*) + L1_{smooth}(P, P^{gt}) \quad (6)$$

3.2 어텐션 분산 제약

각 배치 표본 n 에서, ego 상태의 $C = 6$ 개 채널에 대한 교차어텐션 가중치 $a^{(n)} \in \Delta^{C-1}$ 가 균등분포에서 얼마나 벗어나는지를 아래와 같이 정의한다.

$$d(a^{(n)}) = \frac{1}{C} \sum_{i=1}^C \left| a_i^{(n)} - \frac{1}{C} \right| \quad (7)$$

배치 평균으로 정의하면 아래와 같다.

$$D(\theta) = \frac{1}{N} \sum_n d(a^{(n)}) \quad (8)$$

배치 평균 편차 $D(\theta)$ 가 분산 마진 m 을 넘지 않도록 다음의 같은 부등식 제약을 도입한다.

$$g(\theta) = D(\theta) - m \leq 0 \quad (9)$$

균등 분산일 때 $D = 0$, 채널 하나에 몰리는 극단(원-핫)에 가까울수록 D 가 커지며 본 제약은 ego 상태 어텐션에만 적용된다.

3.3 증강 라그랑지안 방법 기반 업데이트 규칙

본 연구의 제약 최적화 문제는 증강 라그랑지안 방법(ALM)으로 푼다. 먼저, 앞서 정의된 최적화 문제는 다음과 같다.

$$\min_{\theta} \mathcal{L}_{task} \quad s.t. \quad g(\theta) \leq 0 \quad (10)$$

학습의 각 스텝에서 다음의 증강 목적함수를 최소화하여 파라미터 θ 를 갱신한다.

$$\mathcal{L}_{aug}(\theta, \lambda) = \mathcal{L}_{task}(\theta) + \lambda[g(\theta)]_+ + \frac{\rho}{2}[g(\theta)]_+^2 \quad (11)$$

그 후 라그랑주 승수는 아래와 같이 업데이트한다.

$$\lambda \leftarrow \max\{0, \lambda + \rho[g(\theta)]_+\} \quad (12)$$

$$[z]_+ = \max(0, z) \quad (13)$$

제약 항의 계산은 채널 수 $C = 6$ 에 선형인 경량 연산으로 학습 오버헤드는 미미하며, 추론 시에는 제약/승수 계산이 전혀 필요 없어 기본 모델과 동일한 지연 시간을 유지한다. CAR Planner는 동일한 플래너 구조와 손실을 유지하면서, ego 교차어텐션의 분산을 직접 제약해 소수 채널에 대한 솔림(어텐션 붕괴)을 방지하고, 입력을 버리지 않은 채 강건성을 높인다.

3. 실험 결과

훈련은 RTX 4090 기준 20 epoch, 배치 32, Adam(lr=1e-3, wd=1e-4)으로 진행했다. 학습 중에는 state perturbation 을 적용해 현재 ego 상태에 제한된 노이즈를 추가하고, 모든 좌표를 변형된 ego 프레임으로 재정규화했다. 평가는 nuPlan 의 test14-random 과 test14-hard(각 280 개 시나리오, 14 중 유형)에서 수행했다. 여기서 test14-hard 는 시나리오의 난이도가 높게 선별된 분할이다. 지표는 각각 Open Loop Score(OLS), NonReactive-Closed Loop Score(NR-CLS), Reactive-Closed Loop Score(R-CLS)이며 값이 클수록 좋다.

제안 방법과 비교 모델은 아래와 같다.

- Base: ego 의 6 개 상태를 작은 MLP 로 요약 (SDE · ALM 없음).
- PlanTF(SDE): 동일 구조, ego 에 단일 쿼리 상태 어텐션 적용 state dropout encoder 사용
- Ours(CAR Planner): 동일 구조, ego 단일 쿼리 어텐션에 상태채널 분산 제약을 ALM 으로 규제

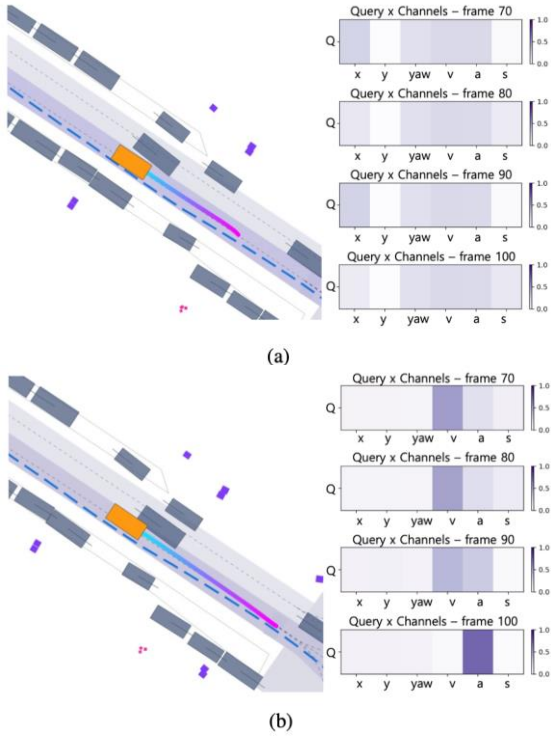


Fig. 2 ego 상태 어텐션 비교. nuPlan test14-hard 분할, R-CLS 시나리오 예시. (a) CAR Planner: (x, y, yaw, v,

a, s) 채널 전반에 어텐션이 프레임 전개 동안 적절히 분산되면서 선택성이 완만하게 유지되어 안전한 조작을 수행함. (b) PlanTF: 어텐션이 점차 일부 채널(특히 a)에 집중(붕괴)되어 상황 균형이 무너지고 결국 충돌로 이어짐.

Figure 2 는 nuPlan 의 test14-hard 분할에서 발췌한 두 가지 전형적 상황을 통해 PlanTF(SDE)와 CAR Planner(ALM)의 이고 상태 어텐션 거동 차이를 제시한다. 두 경우 모두 아키텍처는 동일하며, 차이는 이고 상태 어텐션 모듈뿐이다. PlanTF 는 단일 쿼리 기반의 상태 어텐션을 확률적 채널 드롭아웃(SDE)으로 학습하되 분산 제어가 없어, 프레임이 진행될수록 어텐션이 소수 채널(특히 가속도 a)에 점차 붕괴되어 (x, y, yaw)와 같은 횡방향 단서에 둔감해지고, 혼잡 · 가림 상황에서 불안정한 조작으로 이어진다. 반면 CAR Planner 는 6 채널 어텐션에 증강 라그랑지안 부등식 제약을 적용해 프레임 전반에 걸쳐 분산되면서도 선택적인 패턴을 유지하고, 그 결과 상황 인지가 균형 있게 확보되어 안전한 주행을 달성한다.

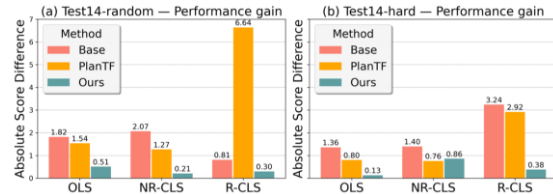


Fig. 3 nuPlan 의 test14-random 및 test14-hard 에서 ego 상태 채널 절제 실험의 격차(state6 - state5)를 나타낸 막대그래프. 값이 작을수록 쇼트컷 의존이 낮음을 의미한다.

Figure 3 은 test14-random 과 test14-hard 에서 채널 감소(state6→state5)시 성능 변화의 절대값(Absolute Score Difference)을 비교한다. 값이 작을수록 채널 감소의 민감도가 낮아 강건하다. (a) test14-random 에서는 Ours 가 OLS 0.51, NR-CLS 0.21, R-CLS 0.30 으로 가장 낮아, PlanTF 의 R-CLS 6.64 처럼 특정 채널 의존이 큰 붕괴가 나타나지 않는다. (b) test14-hard 에서도 Ours 가 OLS 0.13, R-CLS 0.38 로 가장 작고, NR-

CLS 는 0.86 으로 PlanTF(0.76)와 유사한 수준이다. 전체적으로 Ours 는 두 분할에서 모두에서 채널 감소에 따른 성능 저하가 최소로, 쇼트컷 의존이 가장 낮다.

Table 1 test14-hard 평가지표

	Test14-hard		
	OLS	NR-CLS	R-CLS
Base	82.48	65.30	53.11
planTF	83.34	70.03	59.83
CAR Planner	86.31	69.48	64.64

Test14-hard 결과는 Table 1 에 제시된다. CAR Planner 는 OLS 86.31 로 PlanTF(83.34) 대비 +2.97pt, R-CLS 64.64 로 PlanTF(59.83) 대비 +4.81pt 유의미한 향상을 보였다. NR-CLS 는 69.48 로 PlanTF(70.03)와 유사한 수준이다. 즉, 제안 방법은 까다로운 반응형 페루프 환경에서(reactive closed loop) 더 강한 주행 거동을 보이면서, 전반 성능도 우위(OLS)로 확인된다.

Table 2 test14-hard 의 Reactive-Closed Loop 세부 평가지표

	Collisions	TTC	Drivable	Comfort	Progress	Speed
Base	88.11	81.50	92.64	88.23	72.79	98.02
planTF	85.84	80.88	92.64	93.01	84.55	97.01
CAR Planner	90.63	85.49	94.02	98.16	84.28	98.22

Test14-hard 의 R-CLS 세부 평가지표는 Table2 에서 확인할 수 있다. CAR Planner 가 Collision 90.63, TTC 85.49, Drivable 94.02, Comfort 98.16, Speed 98.22 로 더 나은 성능을 보인다. 이는 제안 방법이 기존 방법 대비 더 안전한 주행을 추구하는 경향이 있다는 것을 알 수 있다.

OLS 의 세부 평가지표는 전 구간의 평균 위치 오차 점수인 Average Displacement Error(ADE), 마지막 시점의 위치 오차 점수인 Final Displacement Error(FDE), 전 구간 평균 heading 오차 점수인 Average Heading Error(AHE), 마지막 시점 heading 오차 점수인 Final Heading Error(FHE), 그리고 목표 궤적을 크게 벗어나는 비율에 대한 점수인 Miss Rate(MR)로 이루어진다.

Table 3 test14-hard 의 Open Loop 세부 평가지표

	ADE	FDE	AHE	FHE	MR
Base	79.65	61.34	93.91	91.05	95.95
planTF	81.93	64.31	93.43	90.64	96.32
CAR Planner	84.99	67.81	94.11	91.55	97.06

Test14-hard 의 OLS 세부 평가지표는 Table3 에서 제시된다. Table2 와 Table3 을 종합하면, CAR Planner 는 경로 정확도와 안전성과 안정성을 동시에 개선함을 보인다. 이는 CAR Planner 가 ego 상태 어텐션에 분산 제약을 걸어, 채널 간 과도한 쏠림을 억제하되 상대적 중요도(우선순위)는 유지하게 만든 결과로 볼 수 있다. 그에 따라 가속 등 일부 단서에 치우친 쇼트컷 의존이 줄고, x,y,yaw 등 6 개의 채널이 우선순위는 남긴 채 균형 있게 활용되어 분포 변화가 크고 복잡한 test14-hard 에서도 경로 정확도와 안전 · 안정성이 함께 개선된다. 아울러, 이 제약은 학습 시에만 작동하므로 추론 비용을 늘리지 않는 점도 실용적 이점이다.

4. 결론

본 논문은 모방학습 기반 플래너의 쇼트컷 학습을 완화하기 위해 제약 기반 어텐션 프레임워크인 CAR Planner 를 제안하였다. 어텐션 분산을 제약 최적화 문제로 정식화하고 증강 라그랑지안 방법 (ALM)으로 집행함으로써, 입력 정보를 버리거나 태스크 손실과 간섭하지 않으면서 어텐션 붕괴를 예방한다.

nuPlan 벤치마크 실험에서 CAR Planner 는 test14-random 과 test14-hard 분할 전반에서 기존 SOTA 대비 동등하거나 우수한 성능을 보였다. 특히 쇼트컷을 완화함을 ego 상태 채널 축소 상황에서의 성능 저하를 크게 줄인 것으로 볼 수 있고, 쇼트컷 의존으로 실패가 빈번한 R-CLS 에서 강건성이 향상되었으며 동시에 안전성과 안정성(Collision, TTC, Drivable, Comfort 등) 측면의 이점과 더 매끄러운 주행을 확인하였다. SDE 를 사용하는 PlanTF 등 최신 기준선과 비교해도 분포 이동 및 OOD 환경에서 일반화가 더 강했다.

한계로는 제약이 현재 ego 교차어텐션에만 적용되며, 평가도 nuPlan 에 한정되었다는 점이 있다. 향후에는 제약을 상호작용·맵 수준 표현으로 확장하고, 더 다양한 데이터셋 및 실세계 테스트베드에서 검증함으로써 모방학습 플래너의 강건성과 확장성을 한층 강화할 필요가 있다.

후 기

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (RS-2025-00554087).

References

- 1) D. A. Pomerleau, "Alvin: An autonomous land vehicle in a neural network," Advances in neural information processing systems, vol. 1, 1988.
- 2) M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang et al., "End to end learning for self-driving cars," arXiv preprint arXiv:1604.07316, 2016.
- 3) M. Zare, P. M. Kebria, A. Khosravi, and S. Nahavandi, "A survey of imitation learning: Algorithms, recent developments, and challenges," IEEE Transactions on Cybernetics, 2024.
- 4) S. Ross, G. Gordon, and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," in Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2011, pp. 627-635.
- 5) M. Bansal, A. Krizhevsky, and A. Ogale, "Chauffeurnet: Learning to drive by imitating the best and synthesizing the worst," arXiv preprint arXiv:1812.03079, 2018.
- 6) R. Geirhos, J.-H. Jacobsen, C. Michaelis, R. Zemel, W. Brendel, M. Bethge, and F. A. Wichmann, "Shortcut learning in deep neural networks," Nature Machine Intelligence, vol. 2, no. 11, pp. 665-673, 2020.
- 7) C.-C. Chuang, D. Yang, C. Wen, and Y. Gao, "Resolving copycat problems in visual imitation learning via residual action prediction," in European Conference on Computer Vision. Springer, 2022, pp. 392-409.
- 8) H. Caesar, J. Kabzan, K. S. Tan, W. K. Fong, E. Wolff, A. Lang, L. Fletcher, O. Beijbom, and S. Omari, "nuplan: A closed-loop mlbased planning benchmark for autonomous vehicles," arXiv preprint arXiv:2106.11810, 2021.
- 9) J. Cheng, Y. Chen, X. Mei, B. Yang, B. Li, and M. Liu, "Rethinking imitation-based planners

- for autonomous driving,” in 2024 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2024, pp. 14 123–14 130.
- 10) J. Nocedal and S. J. Wright, Numerical optimization. Springer, 2006.
 - 11) D. P. Bertsekas, “Nonlinear programming,” Journal of the Operational Research Society, vol. 48, no. 3, pp. 334–334, 1997.
 - 12) M. Treiber, A. Hennecke, and D. Helbing, “Congested traffic states in empirical observations and microscopic simulations,” Physical review E, vol. 62, no. 2, p. 1805, 2000.
 - 13) D. Dauner, M. Hallgarten, A. Geiger, and K. Chitta, “Parting with misconceptions about learning-based vehicle motion planning,” in Conference on Robot Learning. PMLR, 2023, pp. 1268–1281.
 - 14) O. Scheel, L. Bergamini, M. Wolczyk, B. Osiński, and P. Ondruska, “Urban driver: Learning to drive from real-world demonstrations using policy gradients,” in Conference on Robot Learning. PMLR, 2022, pp. 718–728.
 - 15) M. Hallgarten, M. Stoll, and A. Zell, “From prediction to planning with goal conditioned lane graph traversals,” in 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2023, pp. 951–958.