청계천 레벨 4 자율주행 버스 도입을 위한 SUMO 기반 시뮬레이션 환경 구축



청계로드 (10조)

조민기(팀장) 강동연 박선우 김희영

01
팀 목표

01 팀 목표

청계A01 버스, 레벨 3에서 레벨 4로



사례: 청계천 도로 A01 버스

목표:

- 단계별로 최대한 현실에 가깝게 구현
- 자율주행 레벨 3 → 레벨 4

완주율

≥ 85%

수동개입률

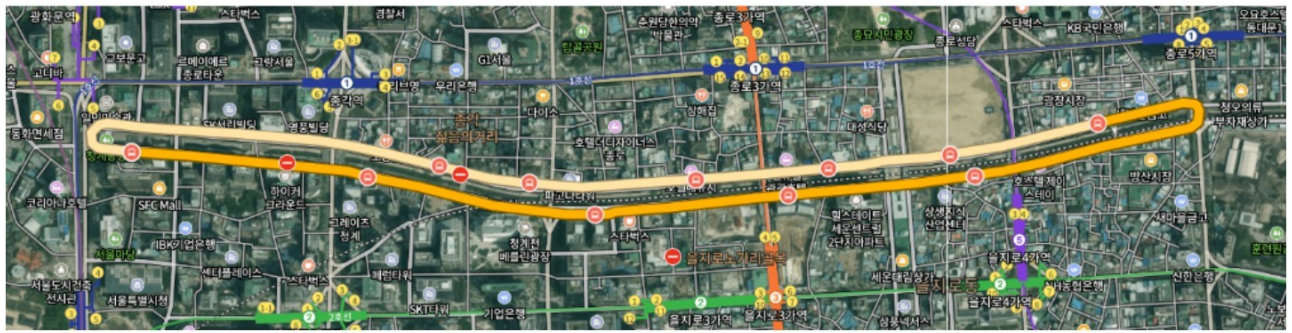
≤ 16%

02
도로 구현

02 도로 구현

실제 노선을 SUMO로

실제 도로 (지도)



SUMO 도로



구축 방법

OSM → netconvert

+ 로드뷰로 차로 수를 직접 보정

차로 수 검증

완료

SUMO에 옮긴 청계A01 노선



버스 노선
2.1 km
청계광장 → 배오개다리

정류장
5곳

신호 교차로
6곳

어린이보호구역
260 m

다섯 단계로 쌓아 올린 환경

처음부터 복잡한 도로 대신, 요소를 하나씩 추가 · 뒤 단계는 앞 단계 요소를 모두 포함



STEP 01
버스
정류장

STEP 02
신호등 /
횡단보도

STEP 03
끼어들기
차량

STEP 04
어린이
보호구역

STEP 05
오토바이
트래픽

버스정류장

정류장

5곳

1. 청계광장
2. 청계1가.광교
3. 청계2가.삼일교(동)
4. 청계3가.관수교(동)
5. 청계4가.배오개다리

정차 성공 조건

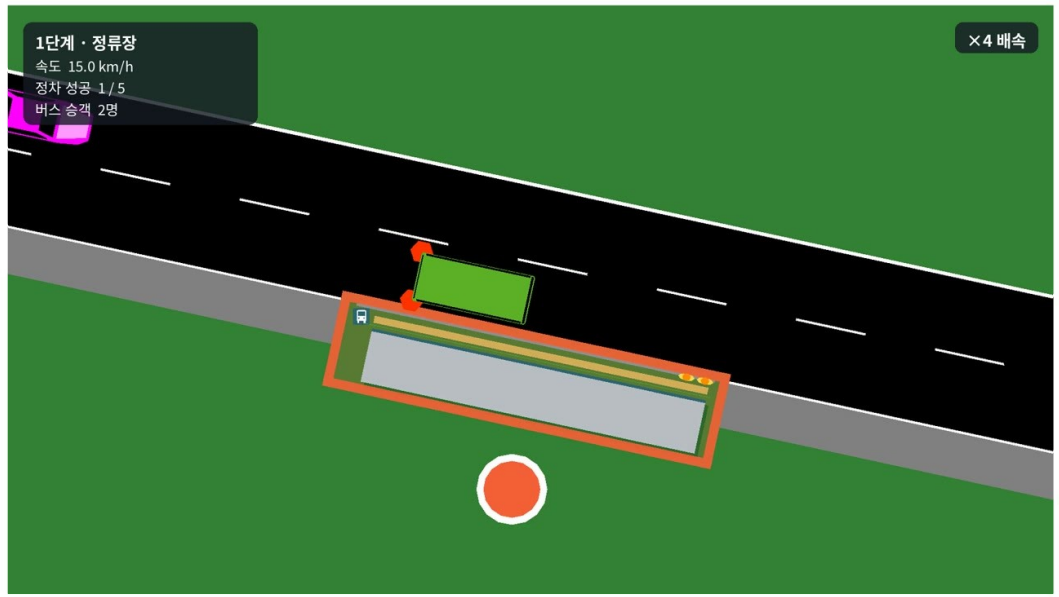
10초 이상

실제 A01 정류장

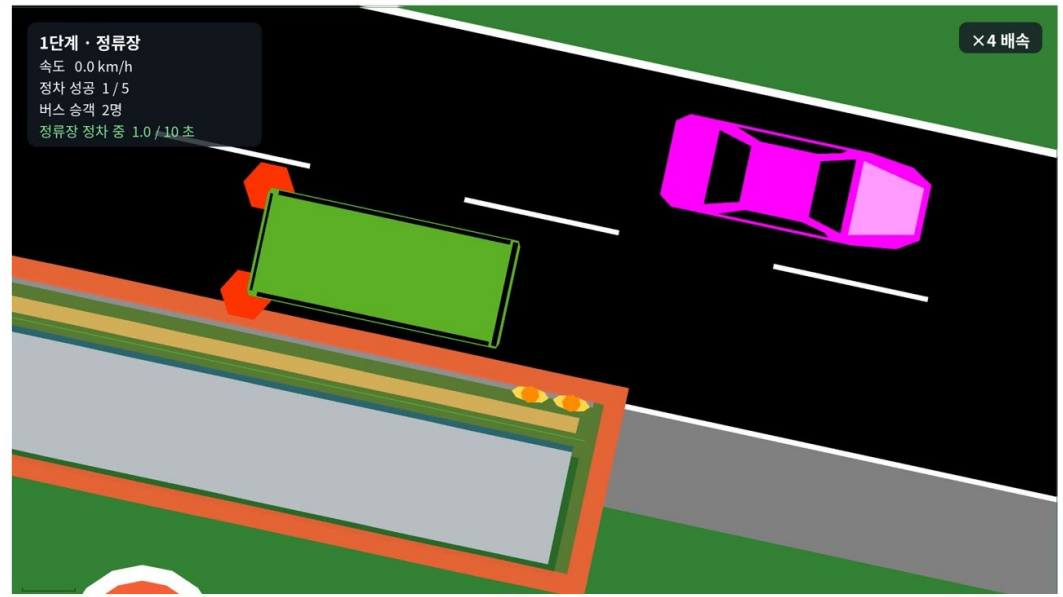


정류장 정차와 승객 탑승

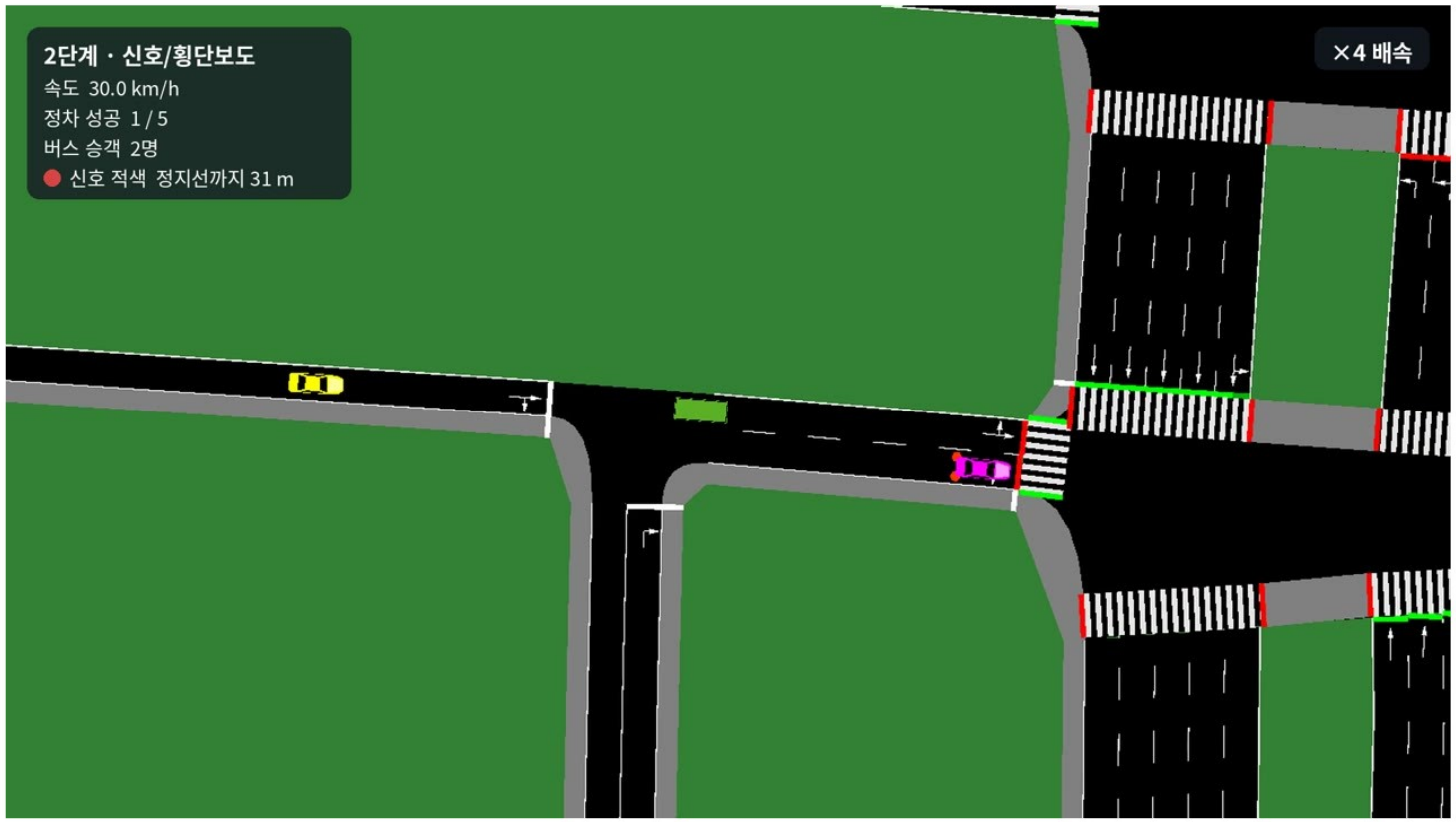
SUMO · 정류장 정차



SUMO · 승객 탑승



신호등 / 횡단보도



실측 신호

6곳

보행자 유입

400명/h

끼어들기 차량



bs_02 앞, 합류 차량 사이로 정류장 차로 진입 · 3단계

흰색 = 옆길에서 들어오는 끼어들기 차량 · 버스는 틈을 기다렸다가 정류장 차로(오른쪽)로 이동

합류 지점

8곳

1. 세종대로
2. 을지로1길
3. 을지로3길
4. 남대문로
5. 옆 연결로
6. 이름 없는 길 3곳

영상: 청계1가.광교 앞

어린이보호구역

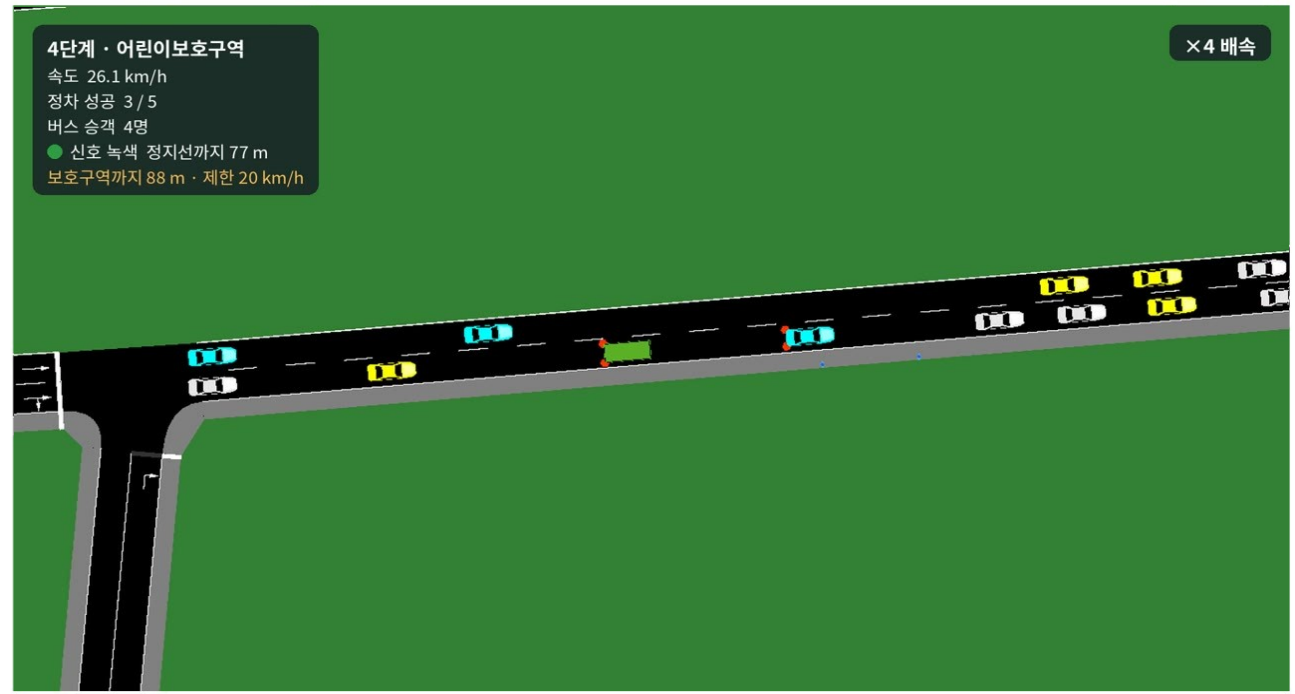
세운상가 앞 약 260 m

일반 구간

30 km/h → **20 km/h**

보호구역

청계천 전체
제한속도



오토바이 트래픽

유입량

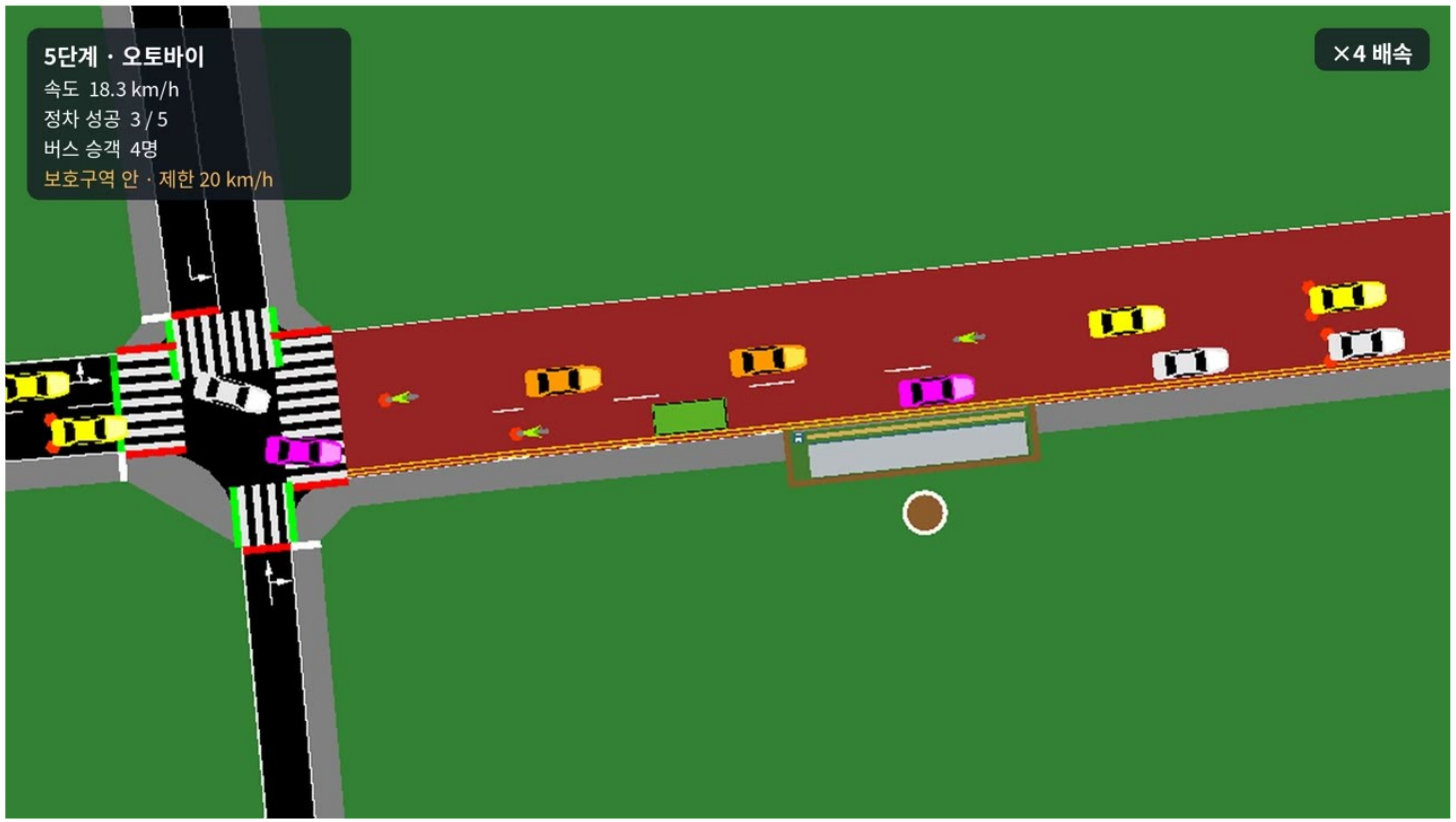
120대/h

차선변경 빈도

1.5배

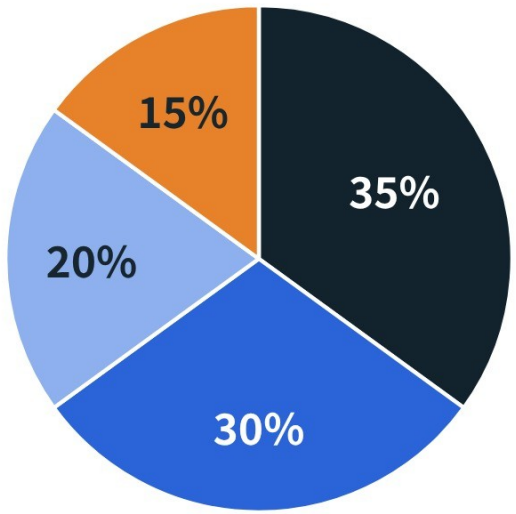
노상 주차 후 복귀

300초 이내



03
일반 차량 구성

본선 승용차: 4가지 운전 성향

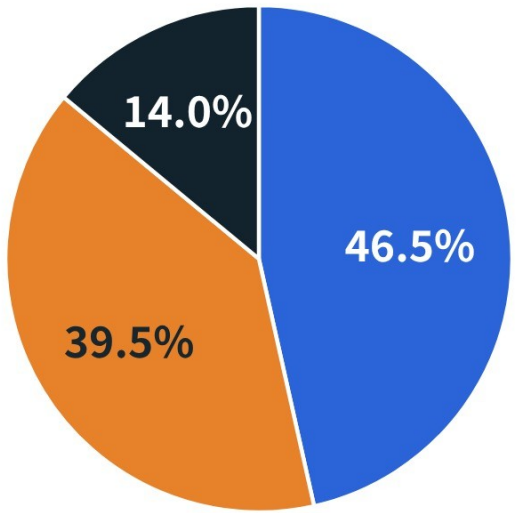


- Krauss 35% (140대)
- IDM 30% (120대)
- EIDM 20% (80대)
- ACC 15% (60대)

= 본선 승용차 400대

모델	비율	어떤 운전자?	운전 스타일
Krauss	35%	딴생각하는 운전자	안전하게 달리다가, 가끔 이유 없이 속도를 줄임
IDM	30%	모범 운전자	앞차와 알맞은 거리를 두고, 부드럽게 빨라지고 느려짐
EIDM	20%	평범한 사람 운전자	앞차가 서면 조금 늦게 알아채고, 가끔 실수도 함
ACC	15%	로봇처럼 꾸준한 운전자	자동 장치가 운전. 앞차와 가장 멀리 떨어져 꾸준히 달림

버스 방향 교통: 시간당 860대



- 본선 승용차 46.5% (400대)
- 끼어들기 39.5% (340대)
- 오토바이 14.0% (120대)

= 시간당 860대

차량	시간당	투입 위치	단계	특징
본선 승용차	400대	본선 입구	1단계~	4가지 운전 성향 혼합
끼어들기 차량	340대	옆길 8곳 (세종대로 60 · 나머지 각 40)	3단계~	더 공격적 & 교차로에서 35% 확률로 양보 없이 진입
오토바이	120대	본선 입구	5단계~	차선변경이 일반 차의 1.5배

04
AV의 관측 · 행동 · 보상

단계별 모델과 학습 방법

규칙이 있는 선생님이 시범을 보이고, 학생은 규칙 없이 관측만 보고 따라 배움

단계		선생님 (시범을 보여 주는 쪽)	학생 (최종 모델)	완주율	충돌률	shield 개입률
1	**정류장**	PPO + 정류장 정차 규칙	BC + DAgger	**99.1%**	0.9%	2.80%
2	**신호**	1단계 선생님 + **적색이면 미리 감속**	BC + DAgger	**90.9%**	2.8%	0.79%
3	**끼어들기**	2단계 선생님 + **넉넉한 틈에서만 차선변경**	BC	**96.3%**	2.7%	0.49%
4	**보호구역**	3단계 선생님 + **보호구역 규칙**	BC	**92.7%**	3.2%	3.66%
5	**오토바이**	4단계 선생님 그대로	BC + KL_PPO	**85.7%**	4.3%	0.31%

1~4단계 1000판, 5단계 400판 기준

총 관측 = 67칸

무엇을 보나	칸 수	버스가 알고 싶은 것
내 속도	1칸	나는 지금 얼마나 빠르지?
주변 차량	30칸	내 차로와 양옆 5개 차로에서 앞차·뒤차는 얼마나 가깝고 빠르지? 앞이 붐비나?
신호	6칸	정지선까지 얼마나 남았지? 지금 무슨 색이지?
경로	18칸	다음 교차로에서 어디로 가야 하지? 어느 차로로 가면 되지?
횡단보도	6칸	앞의 횡단보도 두 곳에 사람이 있나? 나한테 다가오나?
어린이보호구역	2칸	보호구역까지 얼마나 남았지? 몇 km/h까지 달려도 되지?
정류장	3칸	정류장까지 얼마나 남았지? 차로를 몇 개 옮겨야 하지? 얼마나 서 있었지?
차선변경 금지	1칸	지금 차선을 바꿔도 되나?
합계	67칸	버스가 실제로 볼 수 있는 정보만 사용

매 Step 두 가지 행동 선택

행동	값	뜻
가감속	−1	가장 세게 브레이크 (2.5 m/s^2로 감속)
	0	지금 속도 그대로
	+1	가장 세게 가속 (1.0 m/s^2로 가속)
차선변경	−0.5 이하	오른쪽 차로로 옮김
	−0.5 ~ +0.5	지금 차로 유지
	+0.5 이상	왼쪽 차로로 옮김

큰 보상 & 큰 감점

보상	
진행	+10
속도 (빠를수록)	≤ +0.1
완주	+20
정차 성공	+5 / 곳

즉시 종료하는 감점	
충돌	−20
신호등 적색 위반	−20
차로 오류	−20
시간초과	−10

판을 종료하지 않는 감점	
정류장 놓침	−20
보행자가 있는데 횡단보도 진입	−20

05
정량적 주행 성능

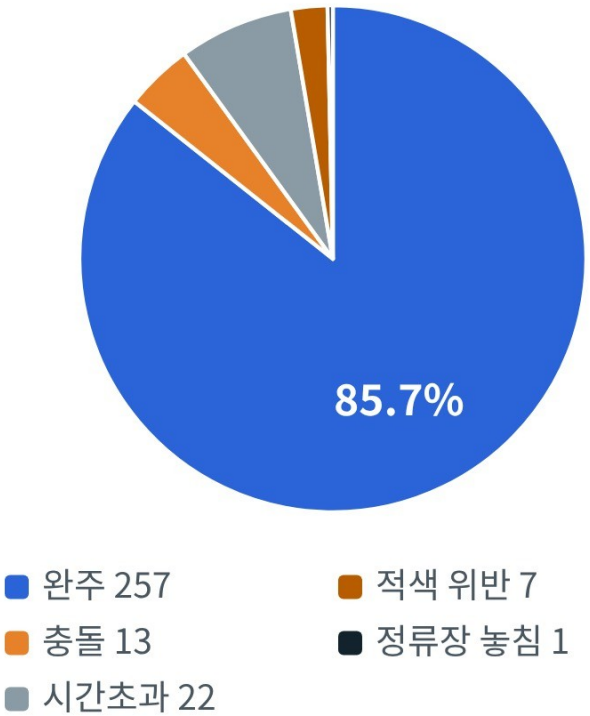
05 정량적 주행 성능

단계별 성능: 충돌률은 모든 단계에서 5% 미만

완주율은 단계가 오를수록 낮아지고 오토바이 단계에서 가장 낮음 · 목표 완주율 95% · 파란 막대는 95% 이상, 주황은 미만

단계	추가 요소	완주율	충돌률	개입률	학습
01	버스정류장	**99.1%**	0.9%	2.80%	BC + DAgger
02	신호등 / 횡단보도	**90.9%**	2.8%	0.79%	BC + DAgger
03	끼어들기 차량	**96.3%**	2.7%	0.49%	BC
04	어린이보호구역	**92.7%**	3.2%	3.66%	BC
05	오토바이	**85.7%**	4.3%	—	BC

5단계 최종 모델 평가 결과: 300판



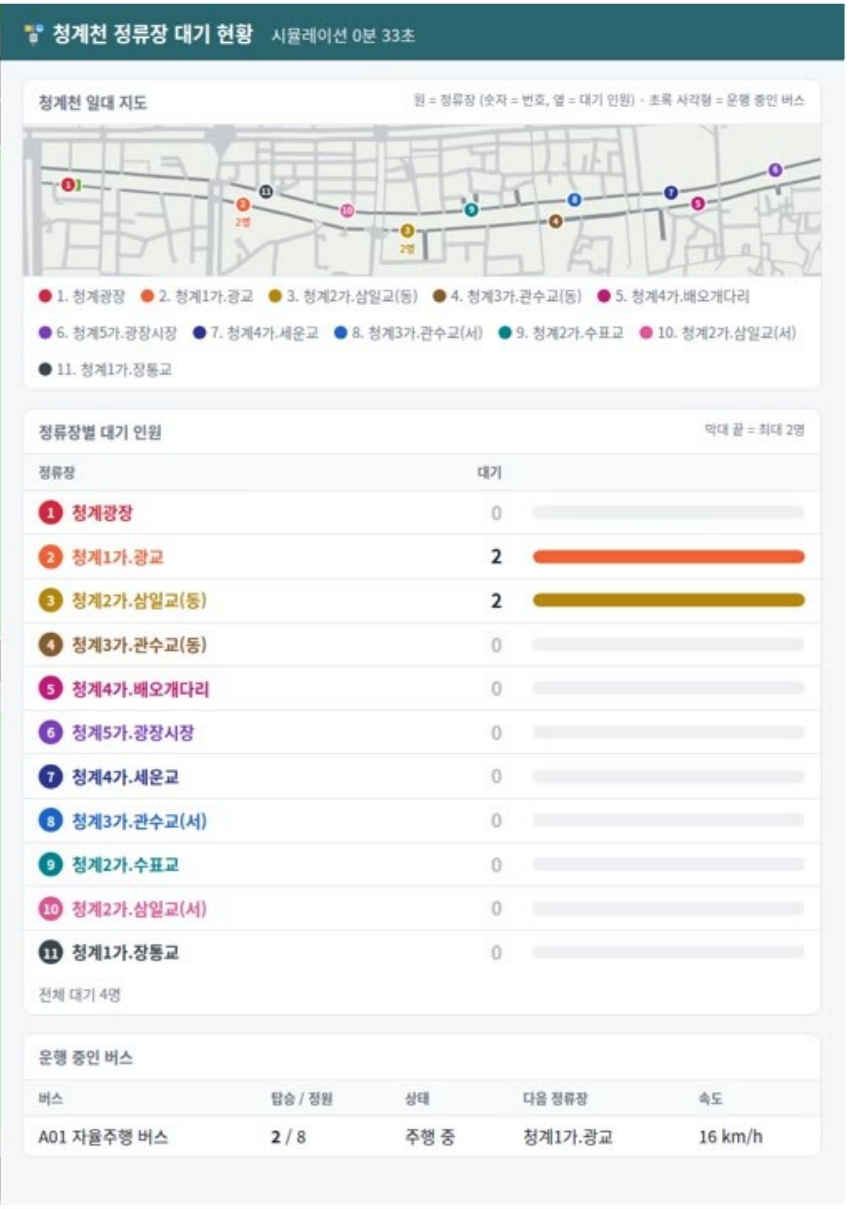
= 전체 300판

지표	정의	목표	결과	판정
완주율	정류장 5곳 모두 정차 + 목적지 도착	≥ 95% (팀 목표)	85.7%	통과
충돌률	충돌로 끝난 판 · 차량 8, 보행자 4, 미확인 1	< 5% (과제 목표)	4.3%	통과
시간초과	1800초 안에 도착하지 못한 판	—	7.3%	—
적색 위반	적색 신호에 정지선을 넘은 판	—	2.3%	—
정류장 놓침	도착했지만 정류장을 건너뛴 판	—	0.3%	—

06

실제 주행 영상





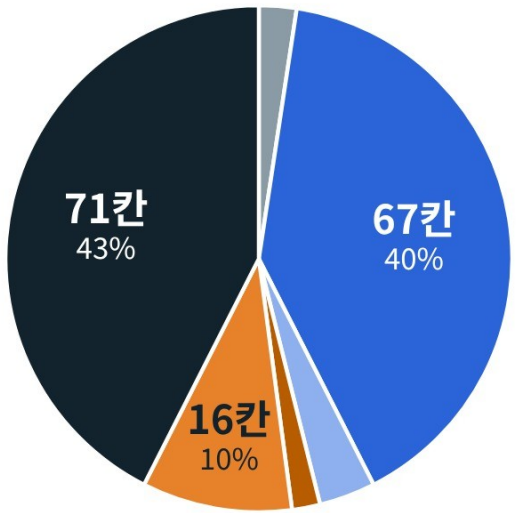
06 실제 주행 영상 · 출발부터 목적지까지 · 5단계 환경 (×15 배속)

07

수집 데이터

07 수집 데이터

매 step마다 167칸을 수집



식별 4 · 보상·종료 3
관측 67 · 사건 16
행동 6 · 3인칭 상태 71

= 167칸

묶음	칸 수	용도	내용
식별	4	판과 시점 구분	판 번호, 시드, step, 경로
관측 obs.*	67	학습 입력	버스가 실제로 보는 정보: 5개 차로 앞뒤 차, 신호, 보행자, 정류장, 보호구역
행동 action*	6	모방학습 정답	가감속, 차선변경 (정책 원본값과 실제 적용값)
보상·종료	3	강화학습 신호	보상, 종료(충돌·도착), 시간초과
사건 event.*	16	평가 지표 계산	충돌, 도착, 적색 통과, 보행자 충돌, shield 개입 4종, 정류장 정차 성공·놓침
3인칭 상태 state.*	71	분석·검증용 (학습에 안 씀)	버스 위치·속도, 신호 7곳의 색과 남은 시간, 정류장 5곳 상태, 승객 수

수집 결과

담당	시드 블록	판 수
1번	5,000,000 ~ 5,001,999	2,000판
2번	5,010,000 ~ 5,011,999	2,000판
3번	5,020,000 ~ 5,021,999	2,000판
4번	5,030,000 ~ 5,031,999	2,000판

학습 시드(1,000~2,999, 900,000~999,999)와 겹치지 않음

수집 조건

모델 : 5단계 BC + KL-PPO
환경 : 5단계
교통 : 버스 방향 860대/h, 반대 방향 400대/h, 보행자 400명/h

Hugging Face
모델 : CheonggyeRoad/cgc-autonomous-driving-bc-ppo · step250k.pt
데이터셋 : datasets/CheonggyeRoad/Dataset

수집한 판

8,000판

transition 총 개수

[11,5xx,xxx]

파일 크기

[약 55 GB]

08
직면한 문제와 해결

문제 1. PPO의 한계

Problem

완주 0%

기본 PPO로 처음부터 학습, 테스트 100판

Approach

선생님 모델
PPO + 규칙으로 학습

↓ 먼저 시범을 보임

BC 모델
관측만 보고 그 행동을 따라 배움

Result

완주 88%

5단계 최종 모델, 테스트 100판
완주 88%
충돌 2%

문제 2. BC 모델이 선생님을 늦게 따라함

선생님 모델  BC 모델

상황	해야 하는 행동	문제	해결
정류장 앞 차선변경	정류장에 다가서기 전에 정류장 쪽 차선으로 옮겨야 함	**타이밍이 늦어 합류하는 차와 부딪힘**	**선생님 모델이 틈이 넉넉할 때만 차선을 바꾸도록 고침**
황색 신호 대응	황색에 알맞게 반응해야 적색에서 제대로 정지	**늦게 반응하면 적색에 진입**	**선생님 모델이 황색 신호에서 더 일찍 서도록 함**

문제 3. 현실 그대로 구현하기 어려웠던 이유

처음 목표: 청계천을 최대한 현실 그대로

무단횡단 | 끼어드는 차 | 갑자기 나타나는 오토바이

불법주정차 | 보행자·자전거·오토바이가 섞인 도로

↓ 이 중 구현한 것

끼어드는 차 | **보행자** | **오토바이**

왜? 모델이 배우지 못함

교수님 피드백

모델 안에 규칙을 넣으면 안 되고,
모델이 스스로 배워서 운전해야 한다

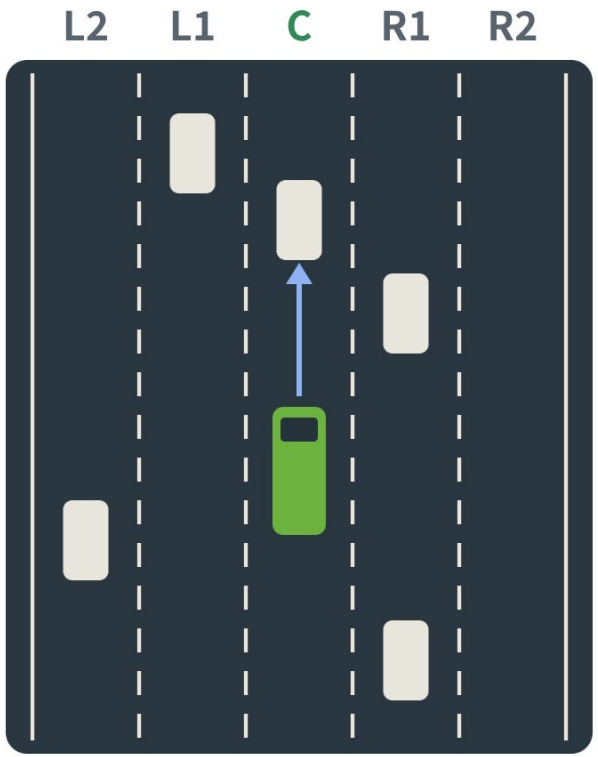
그런데 요소가 많아질수록 모델이 배우지 못함

접근: 특징만 담아 단순하게, 끝까지 어려운 요소는 제거

요소	현실에서는	결정
오토바이	차와 차 사이를 지나다니며 주행	**더 빈번한 차선변경으로 단순화**
자전거	보행자·오토바이와 섞여 주행	**오토바이와 합침**
신호 없는 횡단보도	언제 양보하고 언제 갈지 신호 없이 판단	**제거**
불법주정차	6단계로 구현해 시도	**제거**

감사합니다.

주변 차량 30칸 (5개 차로 × 6가지)



번호	변수	뜻	정규화	없을 때
1–5	lead_gap	앞차까지 거리	÷ 50 m	1
6–10	lead_dv	앞차와 속도 차	÷ 제한속도	1
11–15	follow_gap	뒤차까지 거리 (음수)	÷ 50 m	−1
16–20	follow_dv	뒤차와 속도 차	÷ 제한속도	−1
21–25	lane_end	차로가 끝나는 곳까지	÷ 50 m	−1 차로 없음 · 1 계속
26–30	density	앞 50 m 안 차량 밀도	0 빈 차로 ~ 1 꽉 참	—

- 변수 이름은 obs.(차로).항목
- 차로 = L2 · L1 · C(내 차로) · R1 · R2
- 0번은 내 속도(÷ 13.89 m/s, 30 km/h면 0.6)

신호와 경로 24칸

번호	변수	뜻	정규화 · 값
31	signal.dist	다음 신호 정지선까지 거리	÷ 150 m
32–35	signal.red · yellow · green · stop	신호 색 (stop = 일시정지 후 진행)	하나만 1
36	signal.visible	신호가 보이는가 (버스가 맨 앞일 때만 1)	0 / 1
37–39	turn.left · straight · right	다음 교차로에서 갈 방향 (내비 안내)	0 / 1
40–54	(차로).arrow_left · straight · right	차로별 노면 화살표 (5개 차로 × 3방향)	0 / 1

횡단보도, 보호구역, 정류장 12칸

번호	변수	뜻	정규화 · 값
55–57	crossing0.dist · ped_gap · ped_toward	다음 횡단보도까지 거리 · 보행자와 내 경로 사이 거리 · 다가오는가	÷ 50 m · ÷ 길이 · 0/1
58–60	crossing1.(같은 3가지)	그다음 횡단보도의 같은 값	같음
61	school_zone.dist	보호구역 입구까지 거리 (안이면 0)	÷ 100 m
62	school_zone.limit	보호구역 제한속도	÷ 제한속도
63	bus_stop.dist	다음 정류장 끝까지 거리	÷ 150 m
64	bus_stop.lane_diff	정류장 차로까지 옮길 차로 수 (+ 왼쪽, − 오른쪽)	÷ 2
65	bus_stop.dwell	정류장에서 연속으로 서 있은 시간 (1이면 정차 성공)	÷ 10초
66	lc_zone	지금 차선변경 금지 구간인가	0 / 1

부록 · KL-PPO

KL-PPO란?

BC 모델
선생님 모델을 따라 함 → 선생님을 넘을 수 없음

→

+ PPO 이어 붙이기
더 잘하도록 강화학습

→

완주 0~35%
크게 하락

원인	무슨 일이 있었나	해결 (KL-PPO)
원인 1 **속도 보상으로 편법 학습**	보상표에 "빨리 달릴수록 점수"가 있음	**속도 보상을 0으로** 차선 판단은 BC 그대로 고정
원인 2 **BC를 따르는 힘이 0으로 사라짐**	"BC와 비슷하게 운전하라"는 힘(BC 손실 가중치·KL 벌점)이 학습할수록 줄어 결국 0이 됨	**KL 벌점 세기의 하한 0.01** 끝까지 0이 되지 않게 함