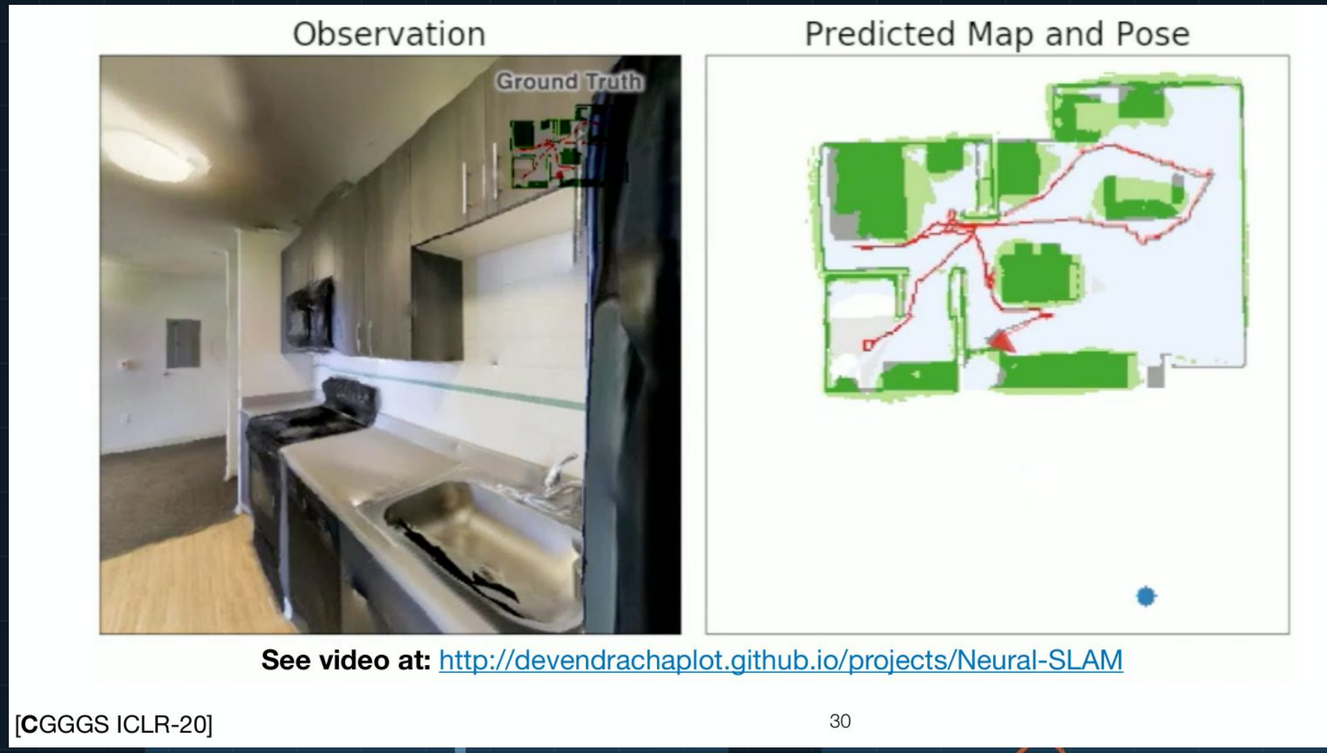
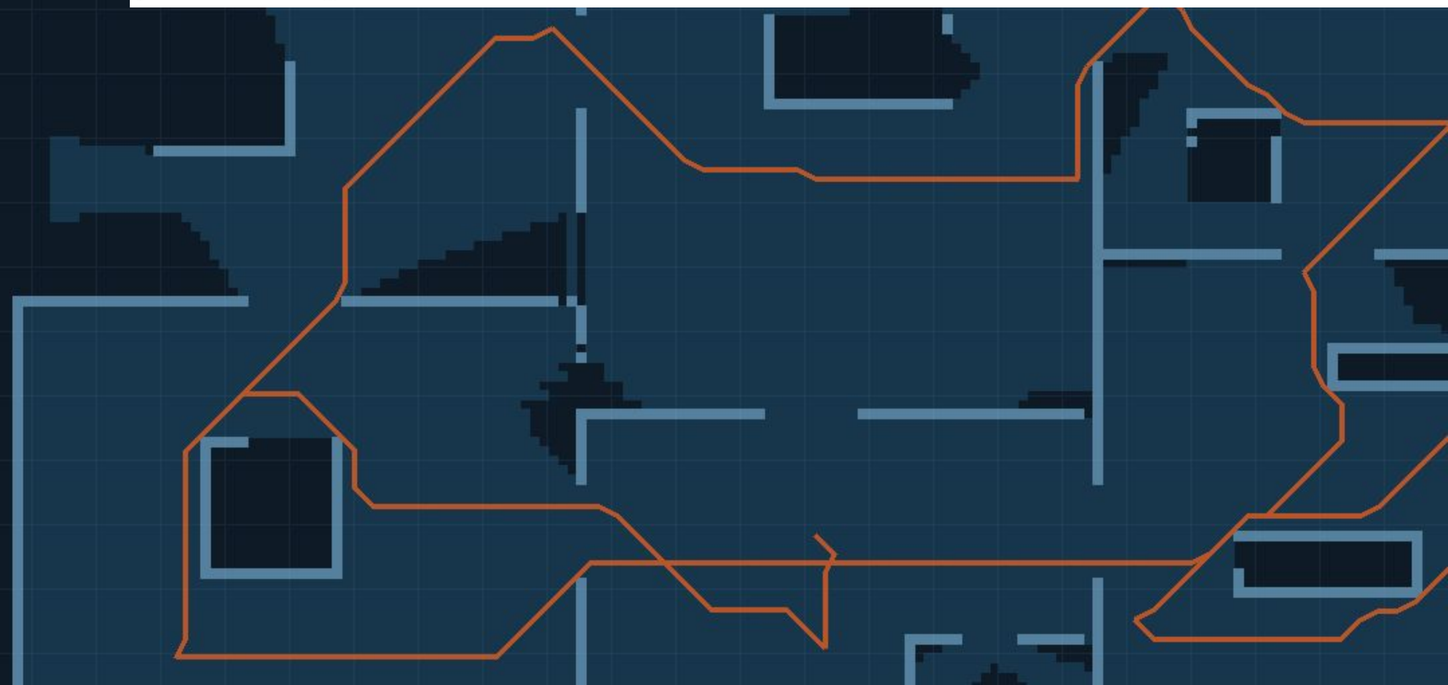
ICLR 2020 · Paper Review

ACTIVE NEURAL SLAM

Learning to Explore using Active Neural SLAM

By 이수경





To a person, a room. To a robot, a photo.

How robots see · 사람에게는 방이 보이고, 로봇에게는 눈높이 사진 한 장이 보인다



사람이 본 거실 (바닥의 로봇 = LoCoBot)



같은 로봇의 카메라가 본 것 (시간순) → 이것만으로 집 전체를 알아내야 한다

카메라 RGB 사진 한 장

바퀴 센서 내 위치 (오차 있음)

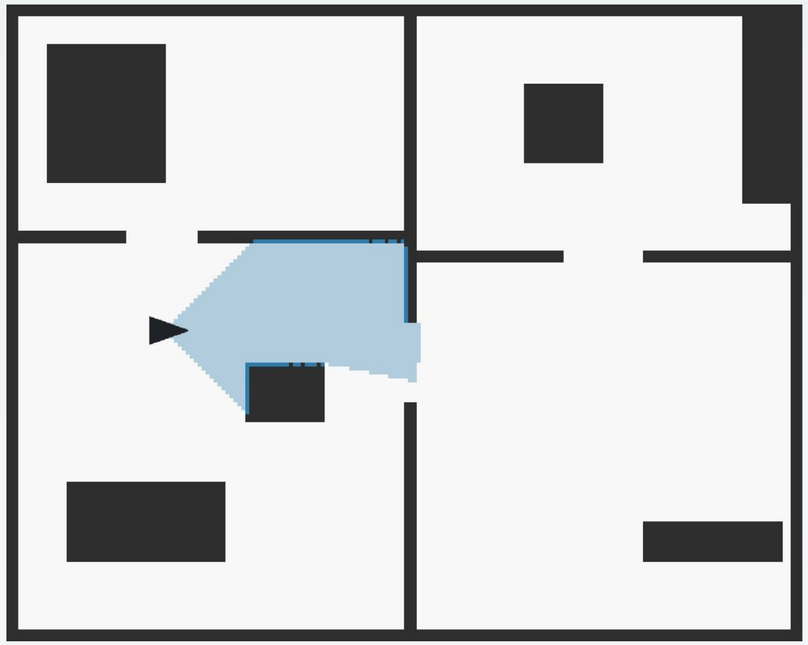
지도 없음 — 처음 보는 집

Chaplot et al., ICLR 2020, Fig 5

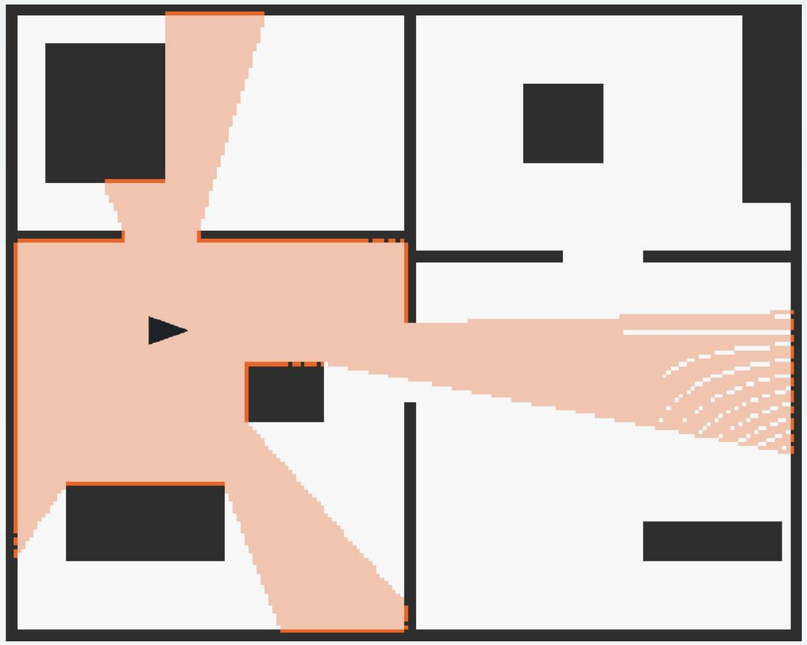
How Robots See

LiDAR measures distance. A camera has to guess it.

RGB 카메라 · 앞쪽 90° · 거리는 사진에서 추측

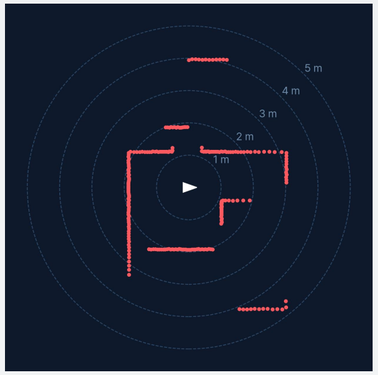


2D LiDAR · 360° · 거리를 레이저로 직접 측정





카메라: 사진



LiDAR: 거리 점

고전 로봇: LiDAR가 360°로 거리를 직접 잰다

ANS: 카메라만 쓴다 → 거리·빈 공간도 사진에서 추측



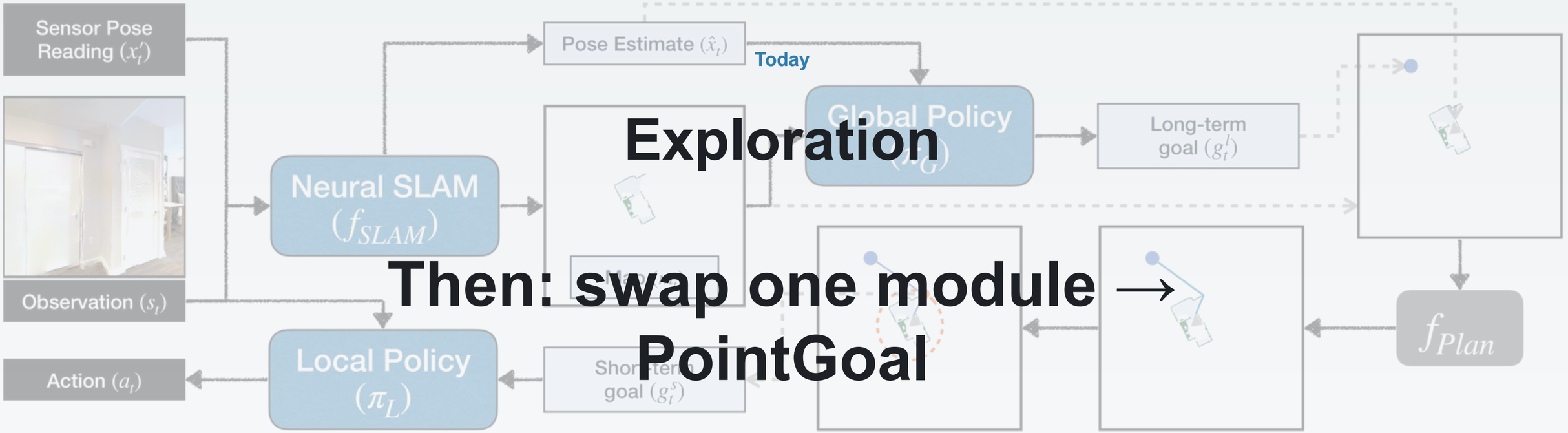
실제 LiDAR가 지도를 그리는 모습 (클릭 재생)

왼쪽 도식·LiDAR 점: 개념도 · RGB: Chaplot, Thesis Defense 2021 · 영상: slam_toolbox (S. Macenski)

Navigation Tasks

Point Goal	Image Goal	Object Goal	Language Goal
 (x, y) *coordinate*		*Chair*	*Go to the blue chair*
		TV	*Go to the largest TV*
		Sofa	*Go to the white sofa*

-> Navigation task converges to exploration



Exploration

Then: swap one module → PointGoal

Chaplot, Thesis Defense 2021

Task

Explore an unseen house in 1,000 steps



RGB camera only

Noisy odometry

No prior map

Score = explored area

Chaplot et al., ICLR 2020

01

Why Modular?



Classical Pipeline

Classical navigation: proven, but LiDAR-bound and myopic

01 **Localization**

Odometry + LiDAR scan matching

02 **Mapping**

Laser hits → occupancy grid

03 **Goal selection**

Nearest frontier

04 **Path planning**

A* · FMM

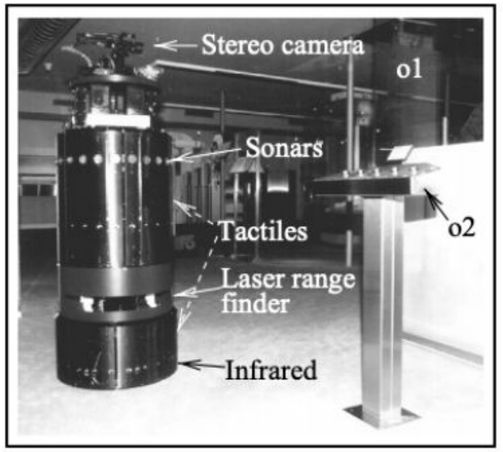
05 **Path following**

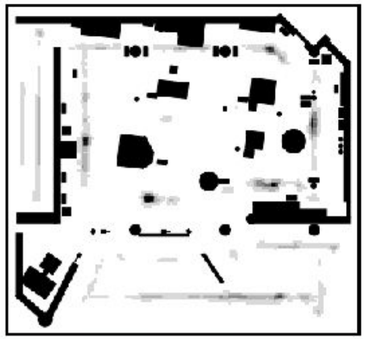
Turn if heading error > 15°

✕ **Geometry needs LiDAR / depth**

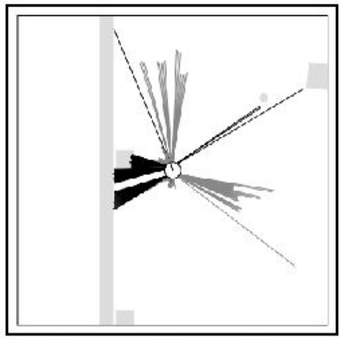
✕ **Nearest frontier → corners first, big picture never**



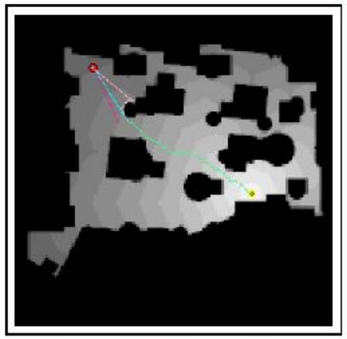




Mapping



Localization



Path Planning

Burgard et al. “The Interactive Museum Tour-Guide Robot”. AAAI-98.

Why modular: keep the map, learn the policy

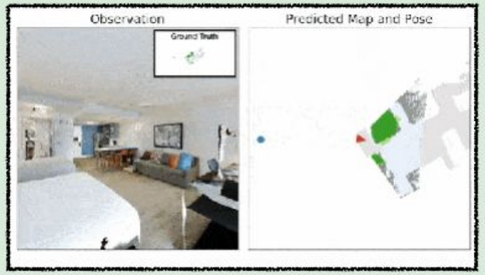
Classical



- Modular
- Explicit Memory/Maps & Planning
- Rule-based policy

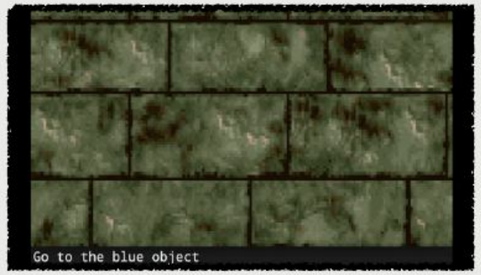
- ⊕ Spatial Understanding
- ⊕ Long-term Memory and Planning
- ⊖ Semantic Understanding
- ⊖ Generalization

Modular Learning



- Learning-based modules
- Explicit structured map representations

End-to-end Learning



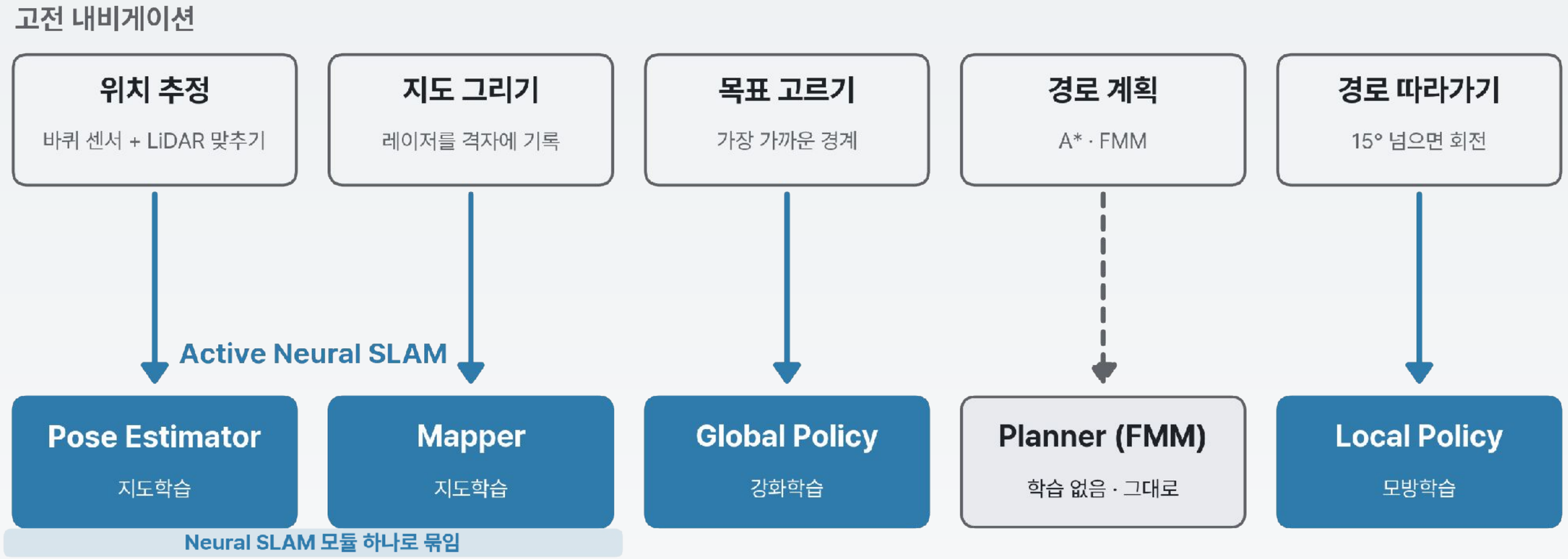
- End-to-end
- Implicit Memory and Planning
- Learned policy

- ⊕ Spatial & Semantic Understanding
- ⊕ Generalization
- ⊖ Long-term Memory and Planning
- ⊖ High sample complexity

학습의 이점 세 가지 (RGB로도 지도 · 센서 오차에 강함 · 집 구조 활용)는 서로 다른 시간 단위에서 일어난다 → 쪼개서 배운다

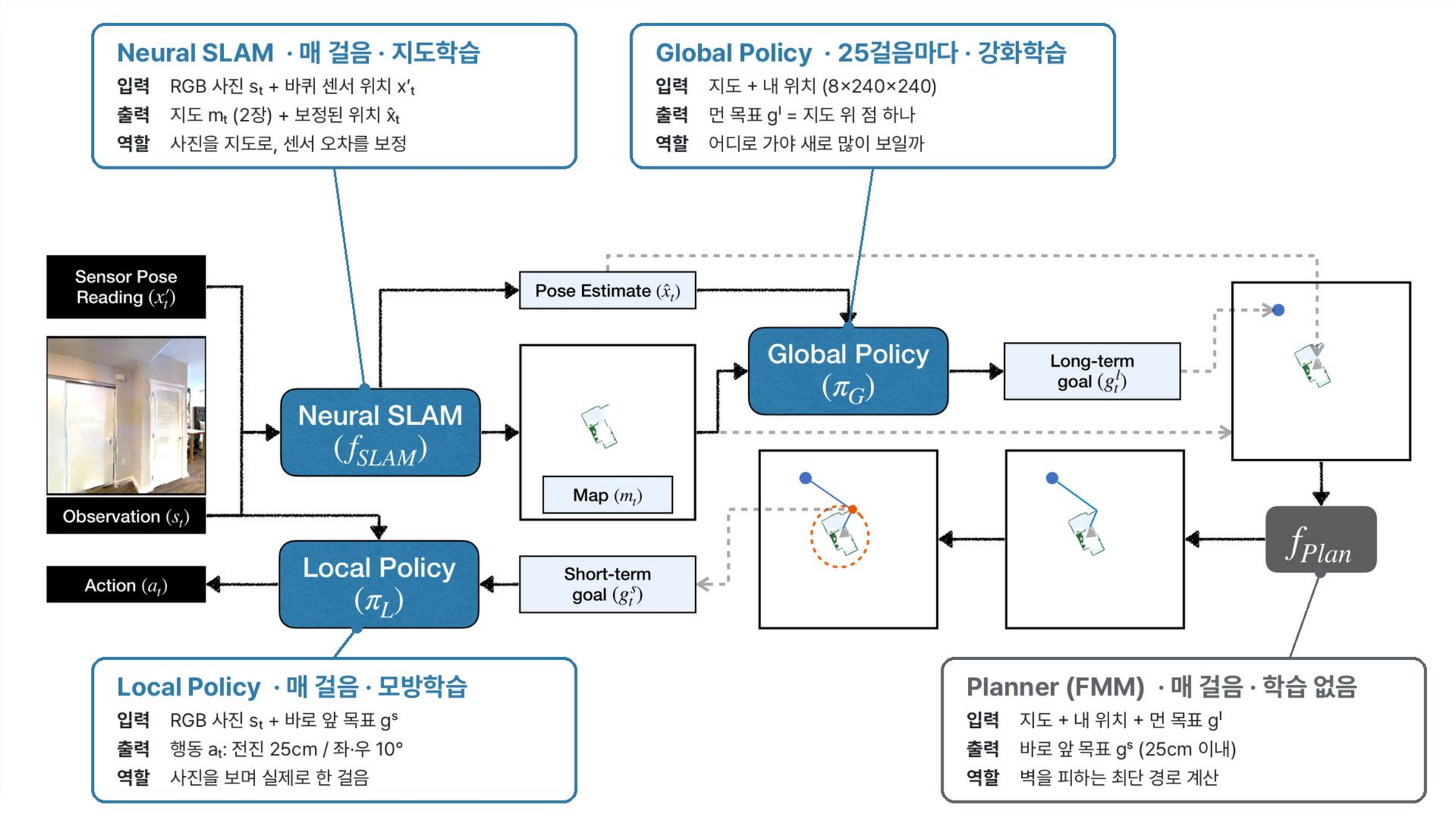
Key Idea

Keep the classical skeleton, swap in learned modules



경로 계획은 이미 정확히 풀리는 문제 → **Planner만 그대로, 나머지 넷은 학습으로 교체**

Four modules talk through a map and a pose



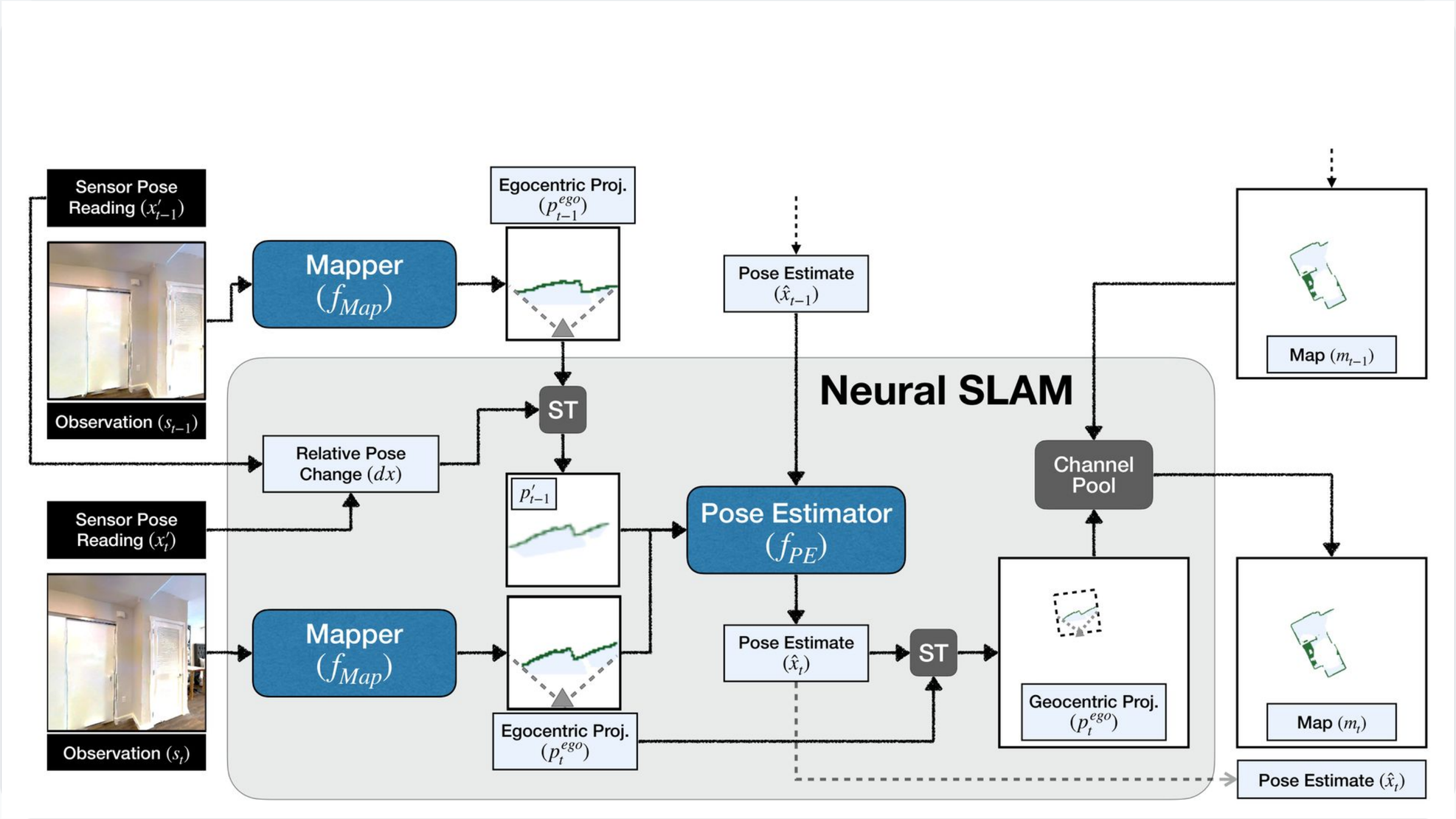
Chaplot, Thesis Defense 2021 (한글 주석 추가)

02

One Cycle, Module by Module



Two photos + odometry → map + corrected pose



Map and pose need each other

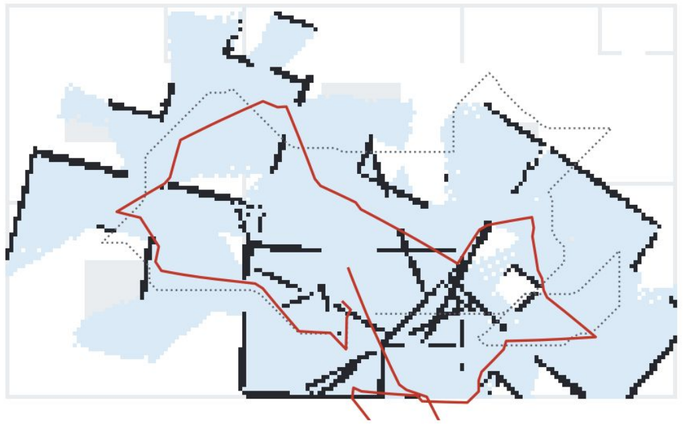
지도가 있어야 위치를 알고,
위치를 알아야 지도를 그린다



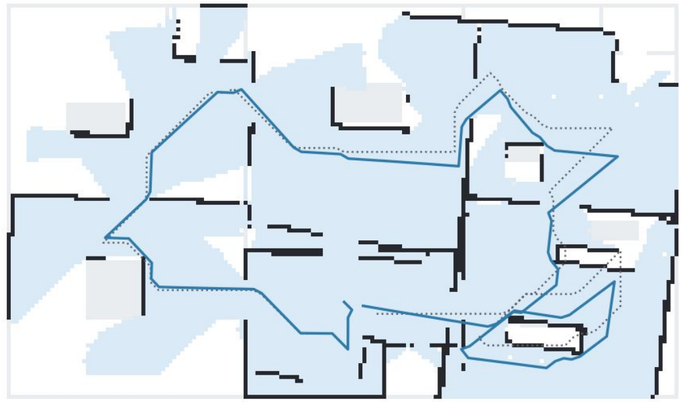
개념 시뮬레이션: 같은 관측을 어떤 위치에 붙이느냐에 따라 지도가 달라진다

바퀴 센서 오차가 걸음마다 조금씩 쌓인다 (회전 약 0.35°/걸음) · 실선 = 추정 경로, 점선 = 실제 경로

센서 값 그대로 (Pose Estimator 없음) 위치 오차 1.32 m



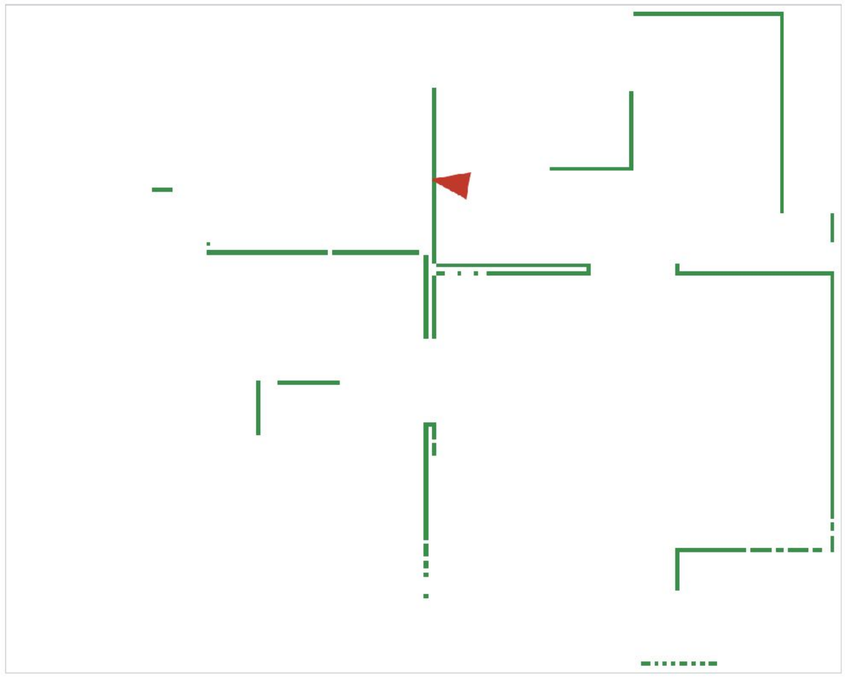
어긋남을 읽어 보정 (Pose Estimator) 위치 오차 0.38 m

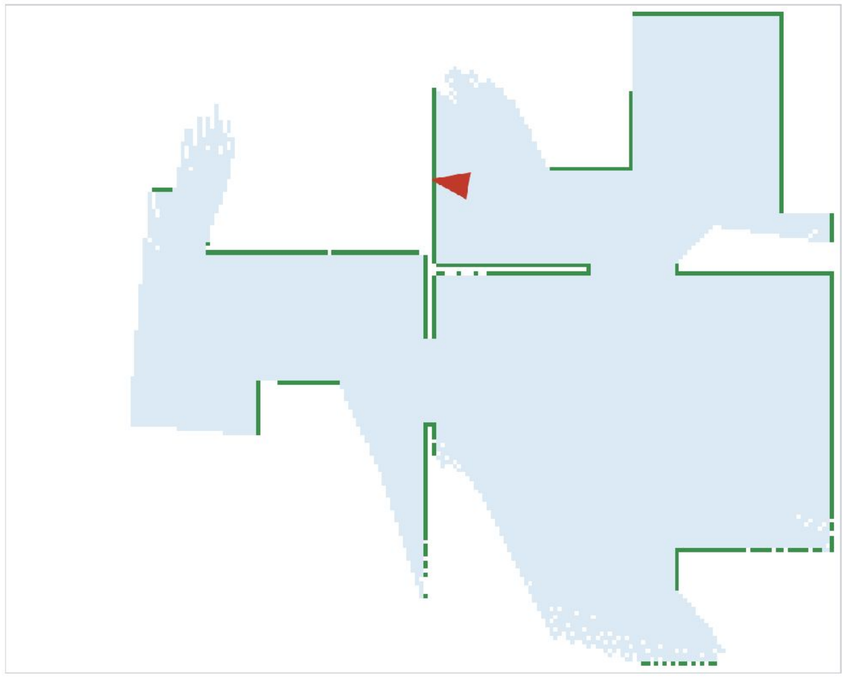


위치가 조금씩 틀리면 → 같은 벽이 다른 곳에 → 지도가 비틀린다

Concept simulation · not the ANS model

LiDAR map: 1 channel. Camera map: 2 channels.





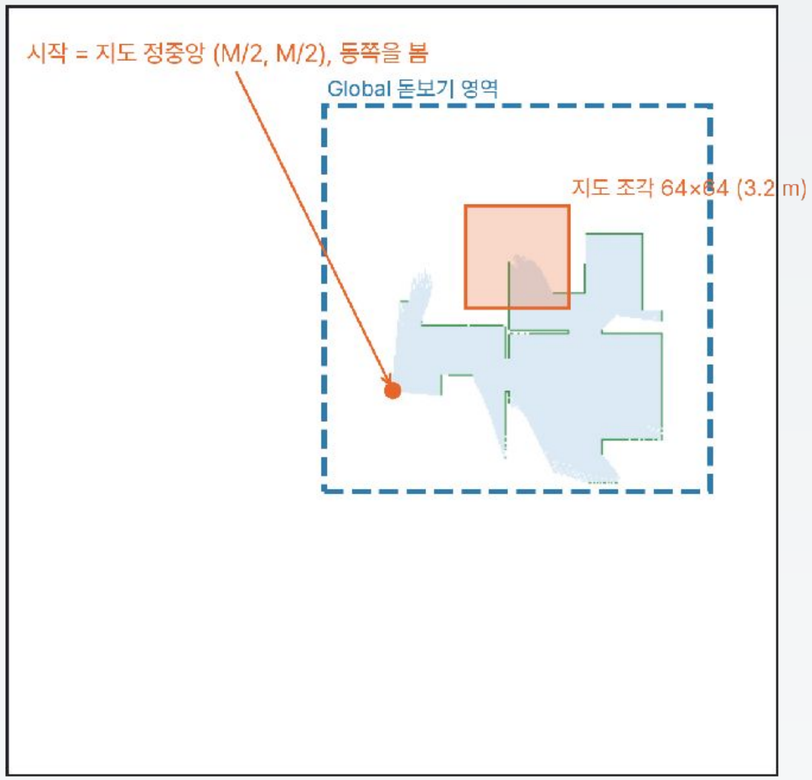
LiDAR: 광선이 지나간 칸 = 빈 곳이라 한 장에 '빈 곳 / 벽 / 모름'을 다 담는다

카메라: '봤나'도 추측해야 하고, 지도를 칸별 max로 쌓기 때문에 장애물 한 장으로는 빈 곳을 못 그린다 → **2장**

개념도 (가상 평면도) · 지도 정의: Chaplot et al., ICLR 2020

The map never grows: fixed size, start at the center

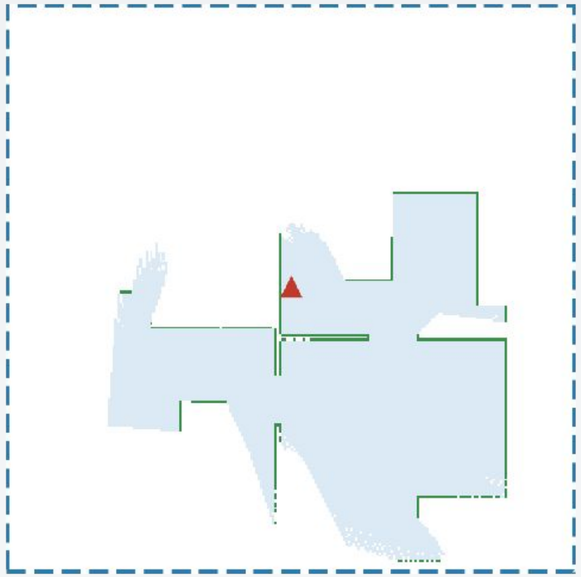
지도는 커지지 않는다 · 크기를 미리 고정



전체 지도 M × M · 처음부터 크게 잡고 시작
학습 480칸 = 24 m · 평가 960칸 = 48 m (칸 5 cm)

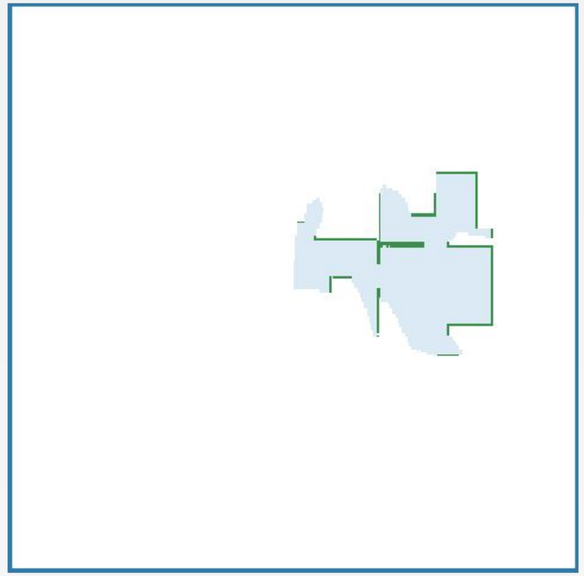
돋보기 · 내 주변 240×240

자르기 (crop) · 12 m · 칸 5 cm 그대로



미니맵 · 전체를 240×240로

max-pool 축소 · 학습 2배 · 평가 4배

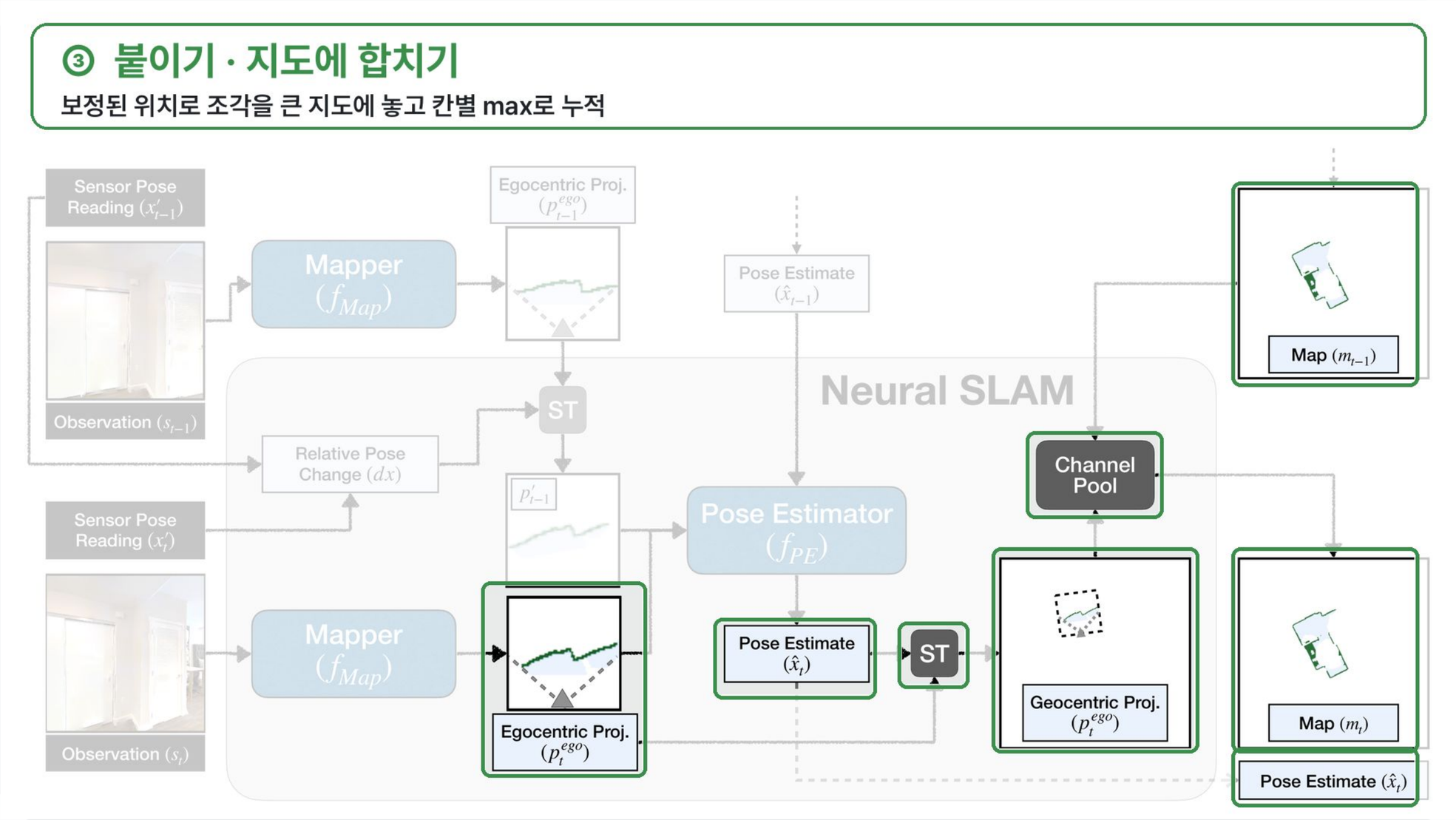


두 장 모두 240×240 → 지도 크기가 달라도 Global 입력은 항상 같은 크기

지도 크기는 처음에 정한다 (학습 24m · 평가 48m) → Global은 항상 **240×240 두 장**만 본다

개념도 · 수치: Chaplot et al., ICLR 2020, Appendix E

Draw → Align → Paste



Chaplot, Thesis Defense 2021 (강조 추가)

Neural SLAM · Training

The simulator knows the answer

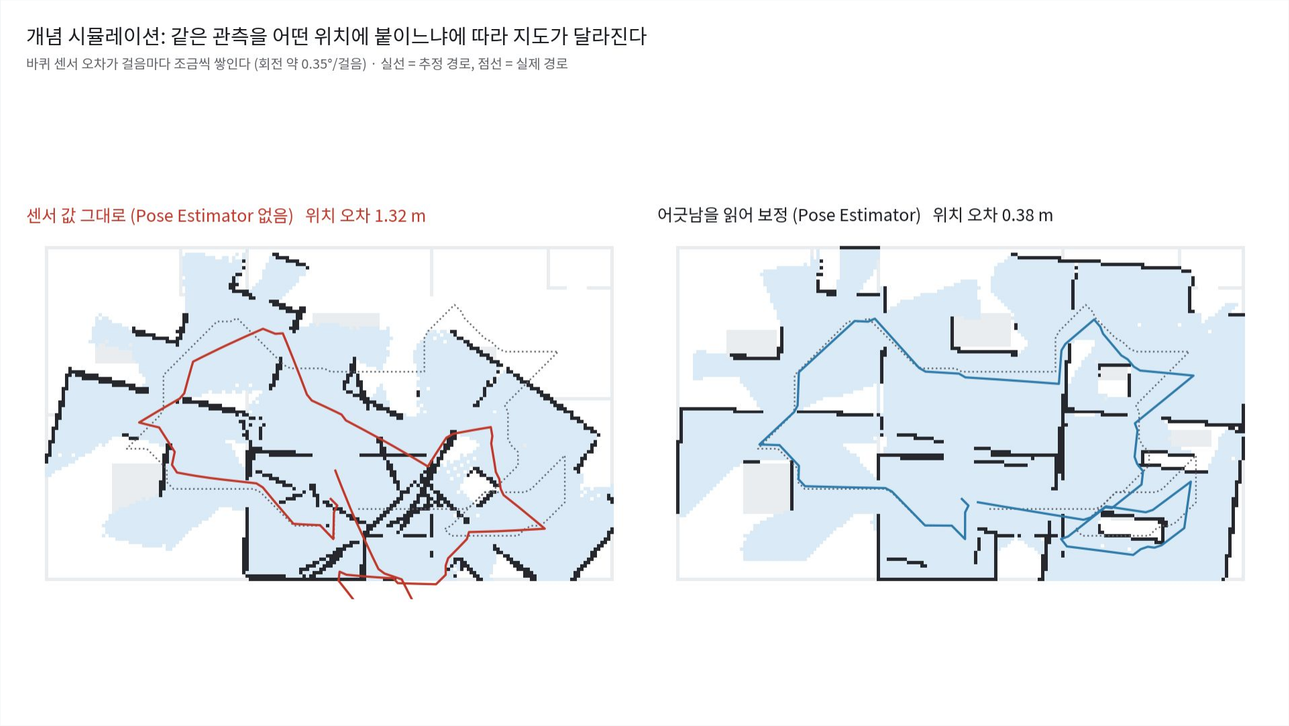
Mapper

정답 = 진짜 depth를 바닥에 투영한 지도 조각

Pose Estimator

정답 = 진짜 이동 − 센서가 말한 이동

센서 오차는 실제 로봇에서 잰 분포로 일부러 넣는다



개념 시뮬레이션 · 실제 ANS 모델 아님

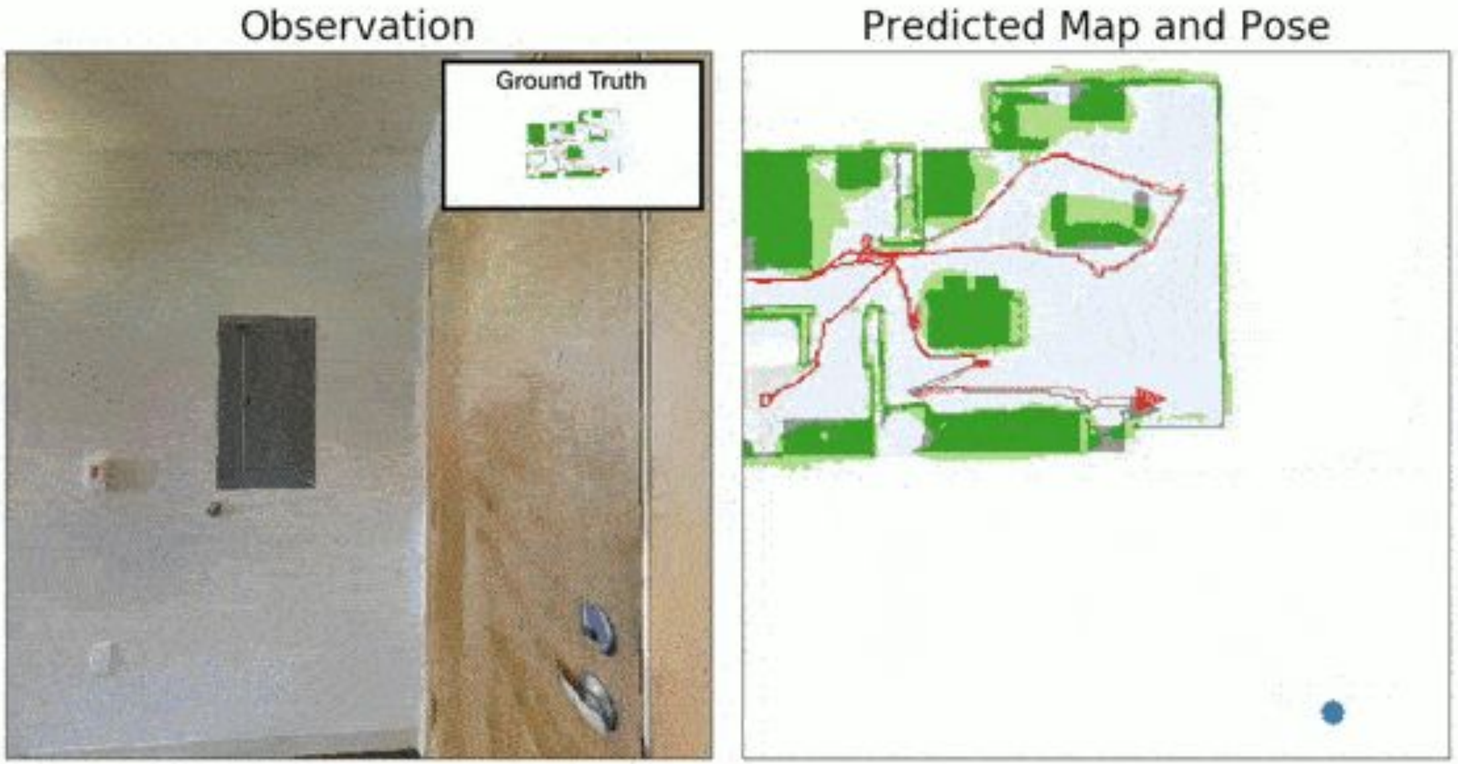
0.862 → 0.771

Pose Estimator를 빼면 · 큰 집 coverage (−9.1%p, 가장 큰 하락)

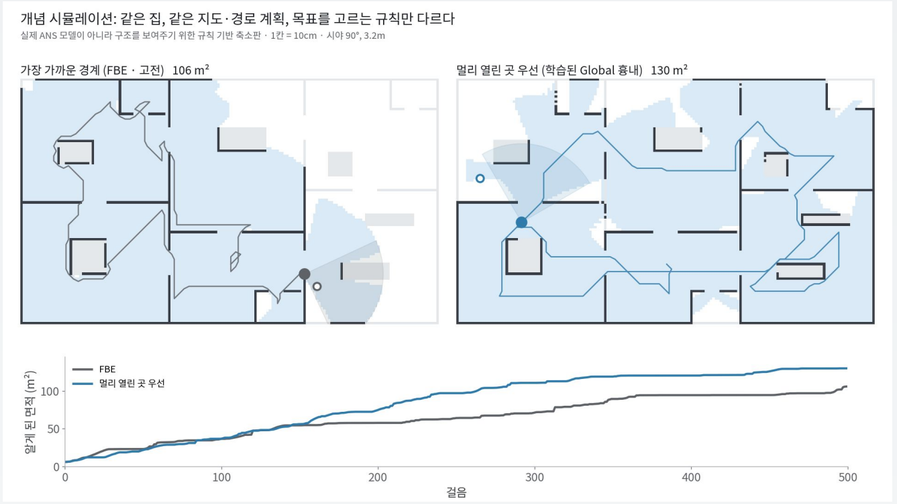
수치: Chaplot et al., ICLR 2020, Table 2

Global Policy

Global policy only drops a pin: where next?



실제 ANS · 파란 점 = Global이 찍은 목표 · 25걸음마다 새로 찍는다



개념 시뮬레이션 · 실제 ANS 모델 아님

목표 고르는 규칙만 다르게 (가까운 경계 vs 멀리 열린 곳)

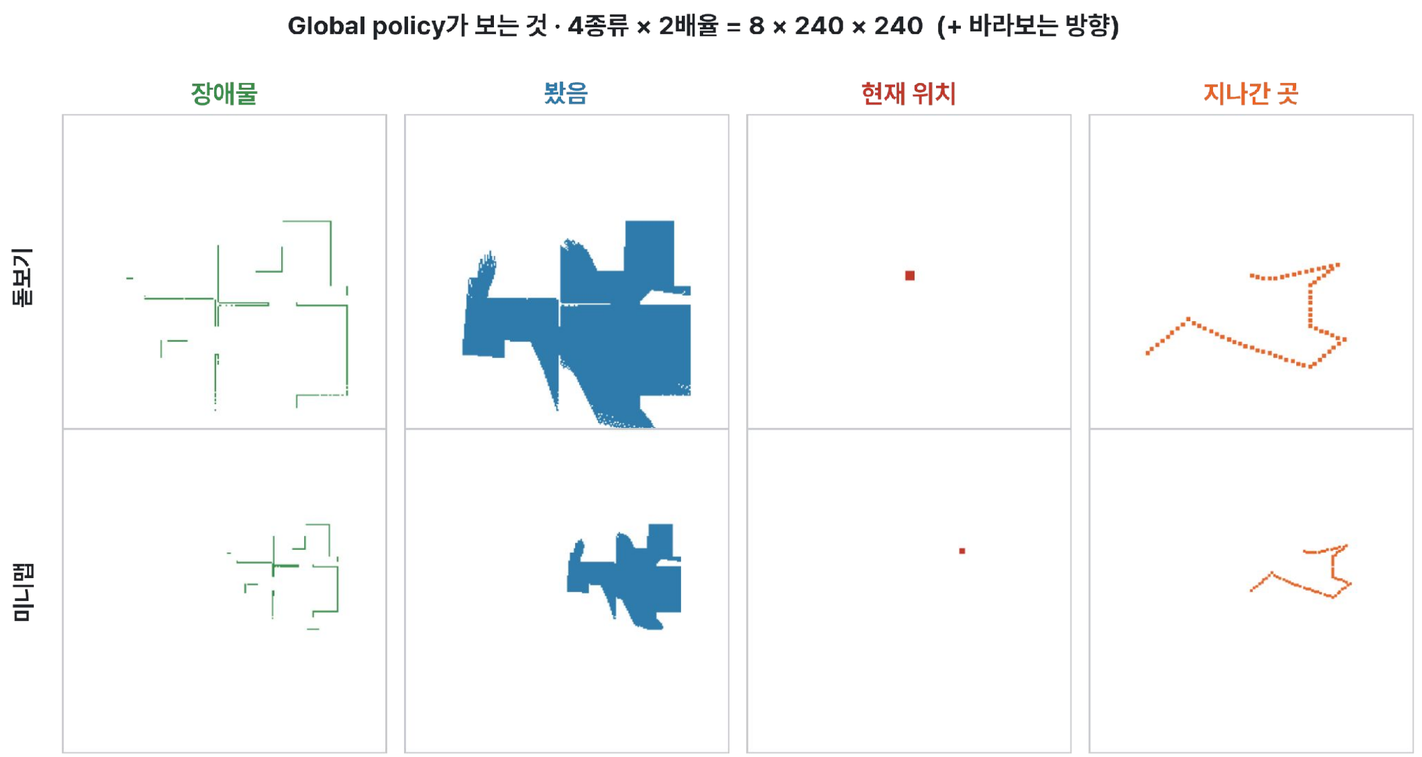
0.862
학습된 Global

0.782
가까운 경계 (FBE)

영상: github.com/devendrachaplot/Neural-SLAM (MIT) · 수치: Table 2, 큰 집

Global Policy · Training

No label for the best goal — learn it from reward



정답이 없다

어디가 최선인지 아무도 모름 → 강화학습 (PPO)

보상

25걸음 동안 새로 알게 된 면적

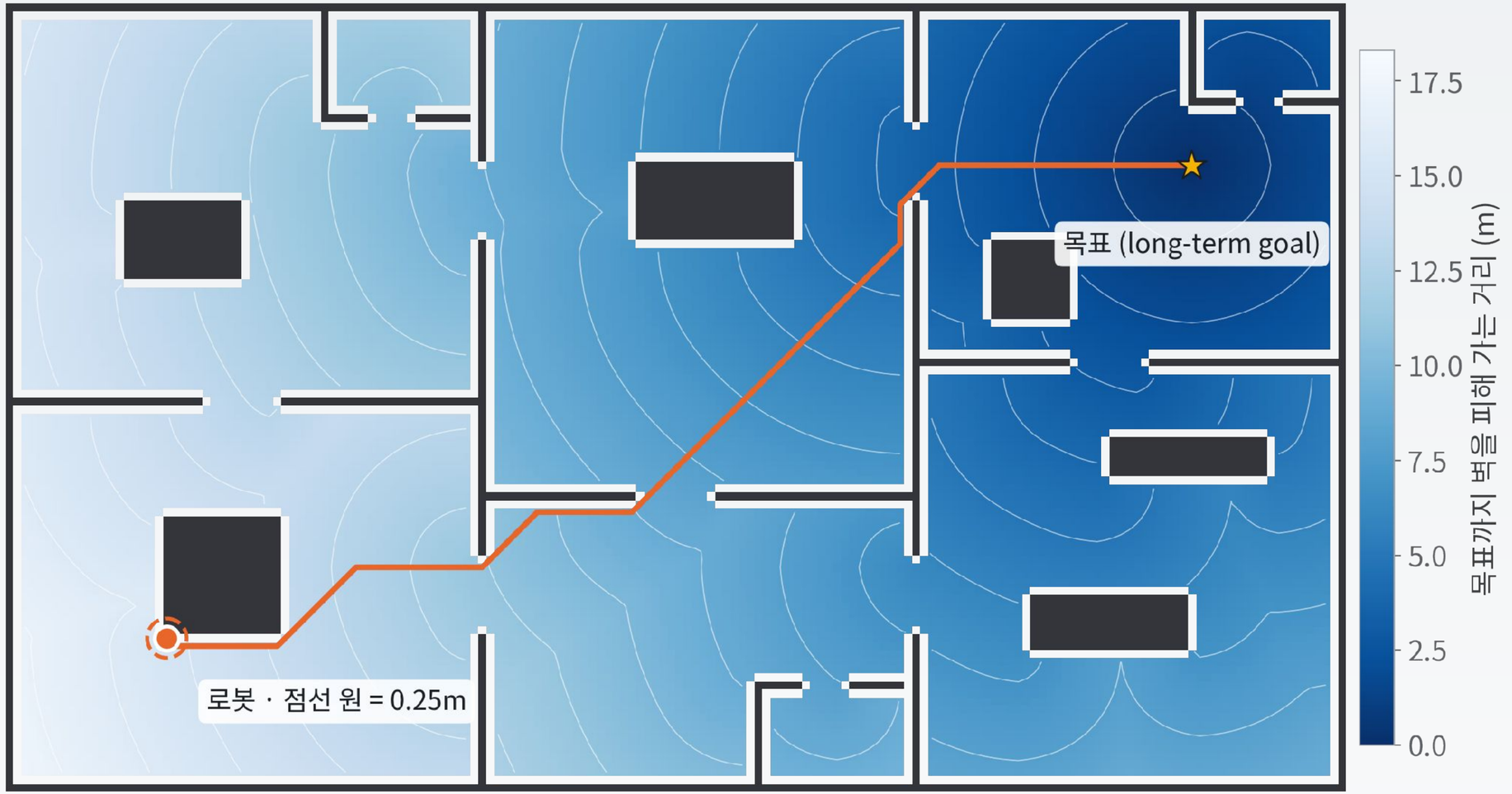
1000 → 40

한 판에 내리는 결정 수

할인율 0.99: 마지막 보상이 첫 결정까지 0.004% → 67% 전달

Planner: no learning, just a distance field

FMM 거리장: 목표에서 물결이 퍼지듯 거리를 매기고, 로봇은 내리막만 따라간다



No learning

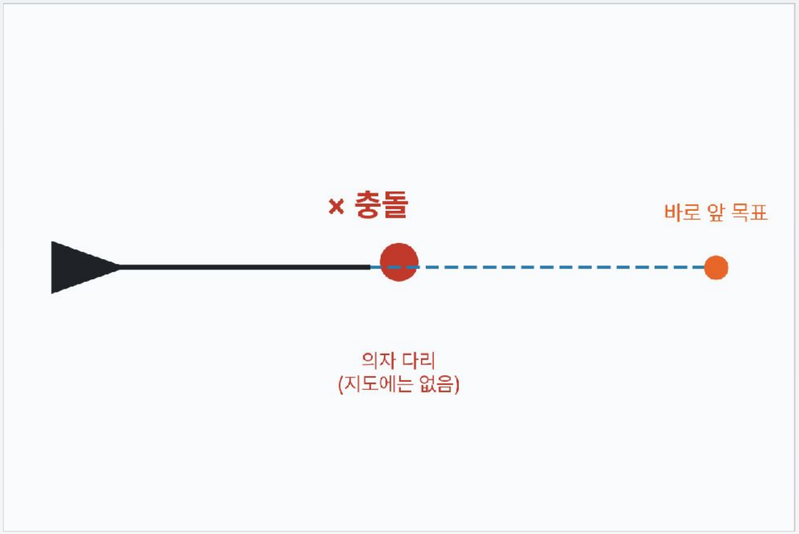
Color = geodesic distance to goal

Unknown = free · replan every step

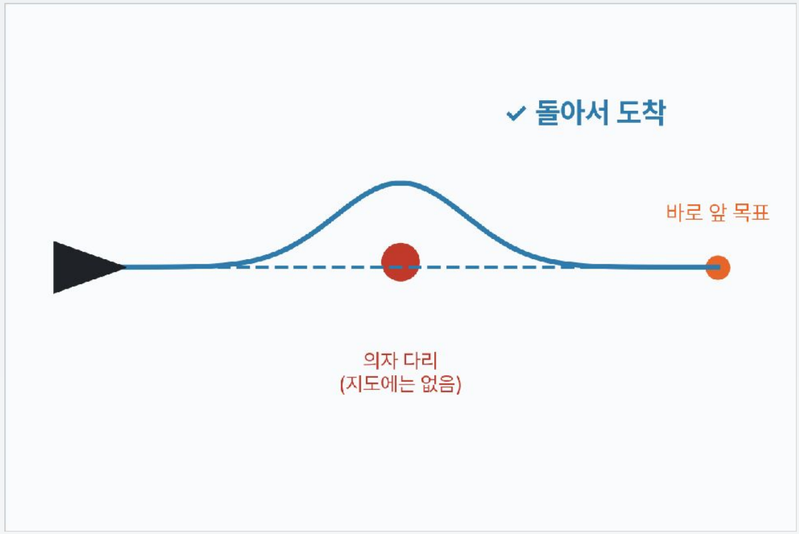
Best point within 25 cm → Local

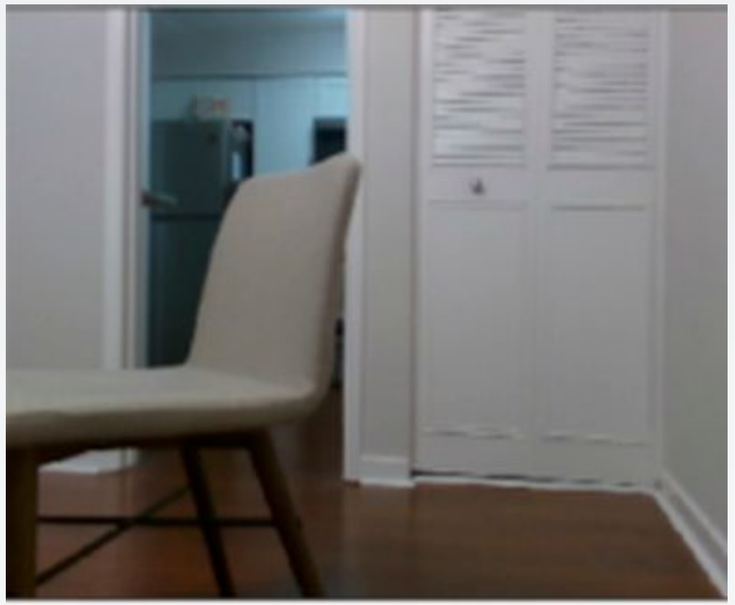
Local policy: dodge what the map missed

15° 규칙 · 지도만 믿고 직진



학습된 Local · 카메라로 보고 피함





실제 로봇 카메라: 지도엔 없던 의자

지도에 없는 작은 장애물을 **카메라로 보고 피한다** — 경로 자체는 못 바꾼다

0.845 → 0.862

15° 규칙 → 학습된 Local · 큰 집 (+1.7%p, 보험)

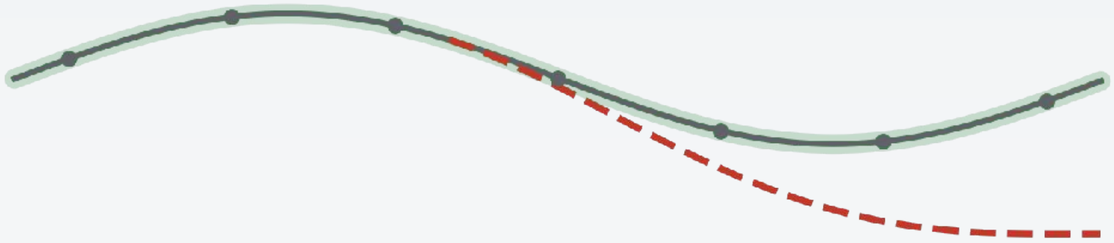
도식: 개념도 · 사진: Chaplot et al., ICLR 2020, Fig 5 · 수치: Table 2

Imitation learning: the student drives, the teacher labels

선생님 = 진짜 지도 위 최단경로 + 15° 규칙 (시뮬레이터 안에만 존재 → 실행 때는 못 씀)

행동 복제 · 선생님이 운전, 학생은 구경

선생님 경로 (진짜 지도 위 최단경로)

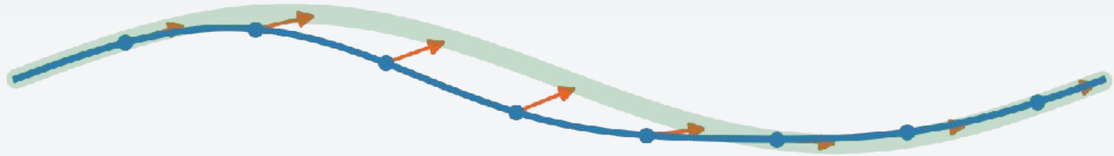


× 처음 보는 상황 → 못 돌아옴

학습 데이터 = 길 위 상황뿐

ANS · 학생이 운전, 선생님이 매 순간 정답

선생님 경로 (진짜 지도 위 최단경로)



✓ 벗어나도 정답을 받아 복구를 배움

주황 화살표 = 그 자리에서 선생님이 고른 행동 (정답 라벨)

정답 라벨이 매 걸음 공짜로 나온다 → 사람 주행 데이터 없이 **사진만으로 선생님 흉내**

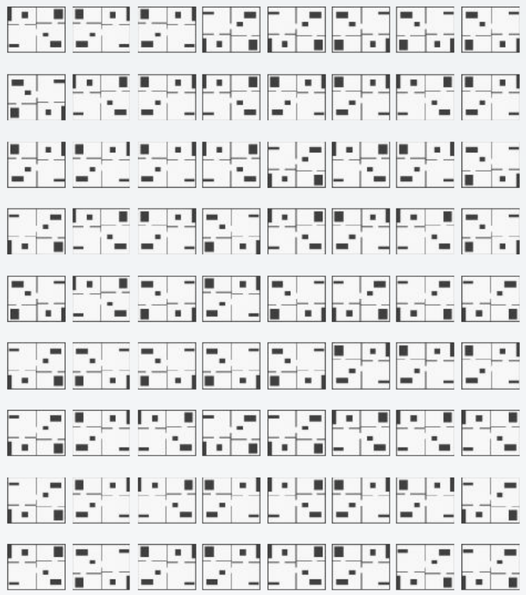
개념도 · Chaplot et al., ICLR 2020

Training · All at once

One trajectory, three separate losses

Gibson 집 72채 · 동시에 탐색

한 판 = 1000걸음 · 학습 총 1000만 프레임



한 줄기 탐색 궤적

사진 · 센서 · 행동

+ 시뮬레이터 정답

(진짜 depth · 위치 · 지도)

모든 부품이 함께 굴러간 결과

Neural SLAM 지도학습

정답 = 진짜 depth 지도 · 진짜 이동

BCE (지도) + MSE (위치) · 버퍼 50만 개 → 72걸음마다 10번 업데이트

Global Policy 강화학습 · PPO

보상 = 25걸음 동안 늘어난 면적

보상 + 가치 추정 (actor-critic) · 한 판(결정 40번)마다 업데이트

Local Policy 모방학습

정답 = 진짜 지도를 아는 선생님의 행동

분류 손실 (전진 / 좌 / 우) · 학생이 운전하며 몇 걸음마다 업데이트

고쳐진 부품이 바로 다음 탐색에 쓰인다

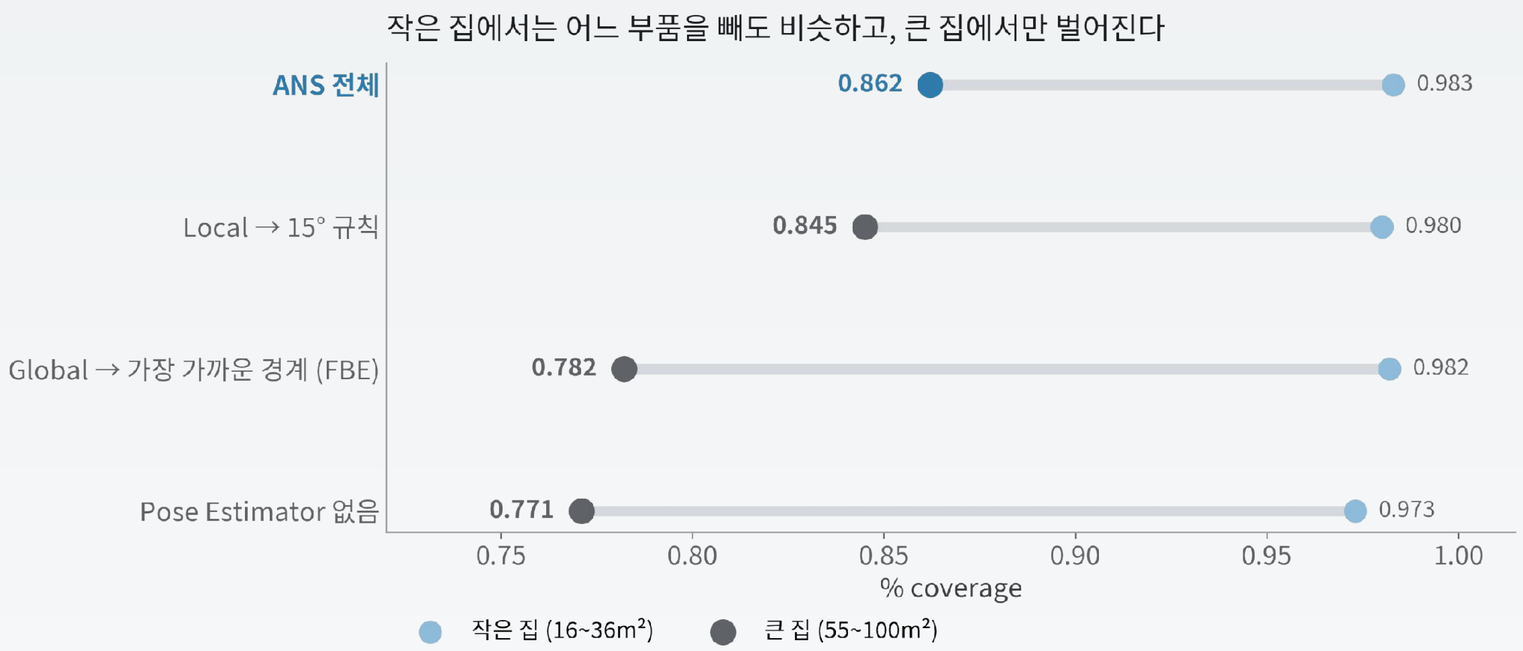
데이터는 함께 만들고, 손실·파라미터는 따로 — 한 부품이 좋아지면 나머지가 받는 입력도 바뀐다

Chaplot et al., ICLR 2020, Sec 5 · Appendix E

Ablation · Gibson Val, large houses

Remove one module at a time — what breaks?

Ablation = 부품 하나를 빼거나 단순한 규칙으로 바꿔서, 성능이 얼마나 떨어지는지 보는 실험



위치 보정(Pose) 제거

−9.1%p 0.862 → 0.771

센서를 그대로 믿으면 지도가 비틀린다

Global → 가까운 경계(FBE)

−8.0%p 0.862 → 0.782

가까운 구석부터 뒤진다

Local → 15° 규칙

−1.7%p 0.862 → 0.845

마지막 한 걸음 보험

작은 집은 무엇을 빼도 ≈ 0.98 — 차이는 큰 집에서만 난다

Chaplot et al., ICLR 2020, Table 2

03

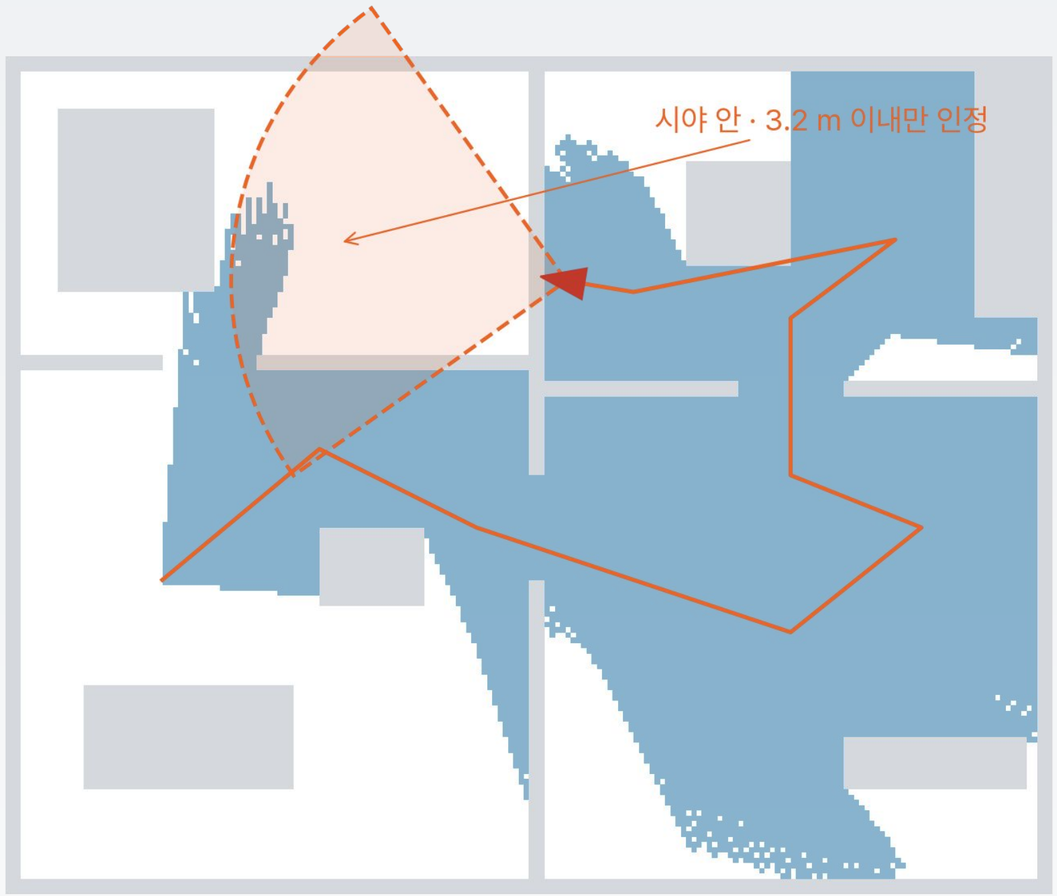
Results



Evaluation

Score = area known to be traversable

Coverage = 지나갈 수 있다고 알게 된 면적



이 예시: 34.3 m² 확인 / 집 전체 63.9 m² → % Coverage = 0.54

Coverage (m²)

지나갈 수 있다고 확인한 면적

시야 안 · 3.2m 이내만 인정

% Coverage

집 전체에서 확인한 비율

한 판 = 1000걸음

처음 보는 집에서만 평가

Gibson 14채 (작은 집 10 · 큰 집 4)

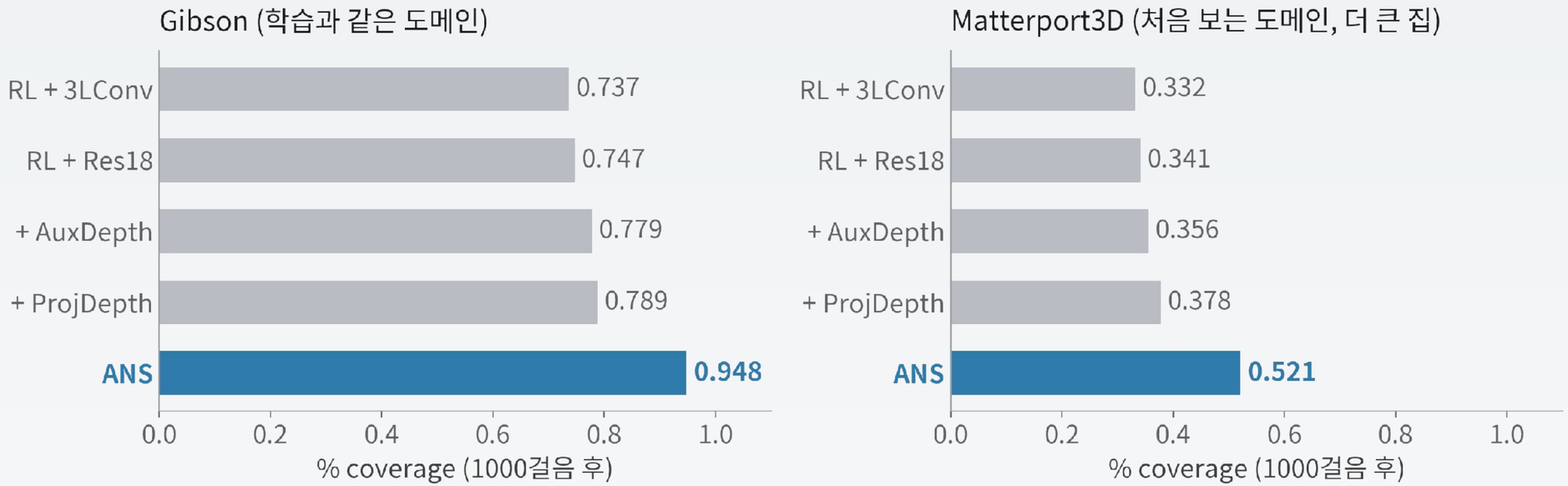
MP3D = 다른 데이터셋, 더 큰 집

도식: 개념도 · 설정: Chaplot et al., ICLR 2020, Sec 5

Main Results

Same data, same reward — ANS beats every baseline

같은 데이터 · 같은 보상 · 같은 PPO로 학습한 end-to-end 4개를 모두 앞선다



0.948

최고 baseline 0.789 · Gibson

73.3 m²

vs 54.8 m² · 처음 보는 도메인 MP3D

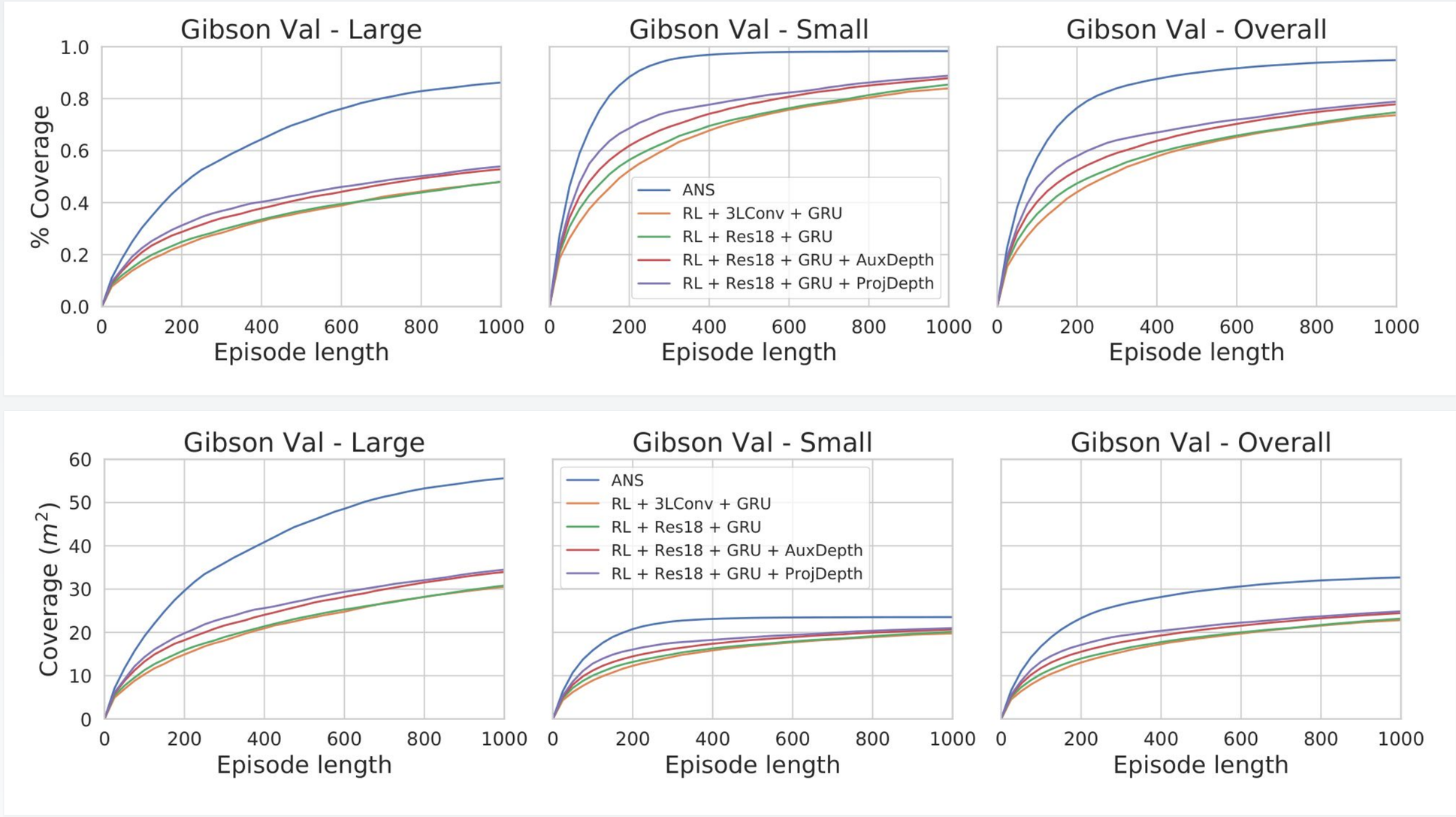
depth를 더 줘도 지도는 안 생긴다

AuxDepth · ProjDepth도 0.79에서 멈춤

Chaplot et al., ICLR 2020, Table 1

Coverage over Time

Small houses saturate early; in large ones the gap keeps growing



Chaplot et al., ICLR 2020, Fig 3

Real Robot

Trained in simulation, deployed on a real robot



Camera height + FOV matched · No fine-tuning

Blue dots = Global goals

Official video ↗

Chaplot et al., ICLR 2020, Fig 5

04

Toward Goal-driven Navigation



Navigation Tasks

Point Goal	Image Goal	Object Goal	Language Goal
 (x, y) *coordinate*		*Chair* *TV* *Sofa*	*Go to the blue chair* *Go to the largest TV* *Go to the white sofa*
Neural SLAM kept	**Global policy** goal = most new area	**Planner** kept	**Local policy** kept

Chaplot, Thesis Defense 2021

Zero-shot Transfer

Harder goals: end-to-end collapses, ANS holds

				Domain Generalization		Goal Generalization			
	Test Setting →	Gibson Val		MP3D Test		Hard-GEDR		Hard-Dist	
Train Task	Method	Succ	SPL	Succ	SPL	Succ	SPL	Succ	SPL
PointGoal	Random	0.027	0.021	0.010	0.010	0.000	0.000	0.000	0.000
	RL + Blind	0.625	0.421	0.136	0.087	0.052	0.020	0.008	0.006
	RL + 3LConv + GRU	0.550	0.406	0.102	0.080	0.072	0.046	0.006	0.006
	RL + Res18 + GRU	0.561	0.422	0.160	0.125	0.176	0.109	0.004	0.003
	RL + Res18 + GRU + AuxDepth	0.640	0.461	0.189	0.143	0.277	0.197	0.013	0.011
	RL + Res18 + GRU + ProjDepth	0.614	0.436	0.134	0.111	0.180	0.129	0.008	0.004
	IL + Res18 + GRU	0.823	0.725	0.365	0.318	0.682	0.558	0.359	0.310
	CMP	0.827	0.730	0.320	0.270	0.670	0.553	0.369	0.318
	ANS	**0.951**	**0.848**	**0.593**	**0.496**	**0.824**	**0.710**	0.662	**0.534**
Exploration	ANS - Task Transfer	0.950	0.846	0.588	0.490	0.821	0.703	**0.665**	0.532

ANS · trained on PointGoal

beats every baseline · 0.951 Succ on Gibson

ANS · trained on Exploration

0.950 Succ · no retraining · almost the same row

Hard-Dist

RL baselines ≤ 0.013 · ANS 0.662 / 0.665

Chaplot et al., ICLR 2020, Table 3

Sample Efficiency · Challenge

75× less data, 1st place in both challenge tracks

ANS · 1M frames

0.789 Succ

Best RL baseline

7500만 프레임 (75×)을 써도 0.678

“ANS achieves > 75× speedup as compared to the best RL baseline.” — Appendix A

CVPR 2019 Habitat Challenge · team “Arnold”

RGB Leaderboard

Rank	Team	SPL
1	Arnold	0.805
1	Mid-level-Features	0.800
3	CHROMA	0.712
4	ARF-RL	0.699
5	MTank	0.260
6	Policy-Police	0.260

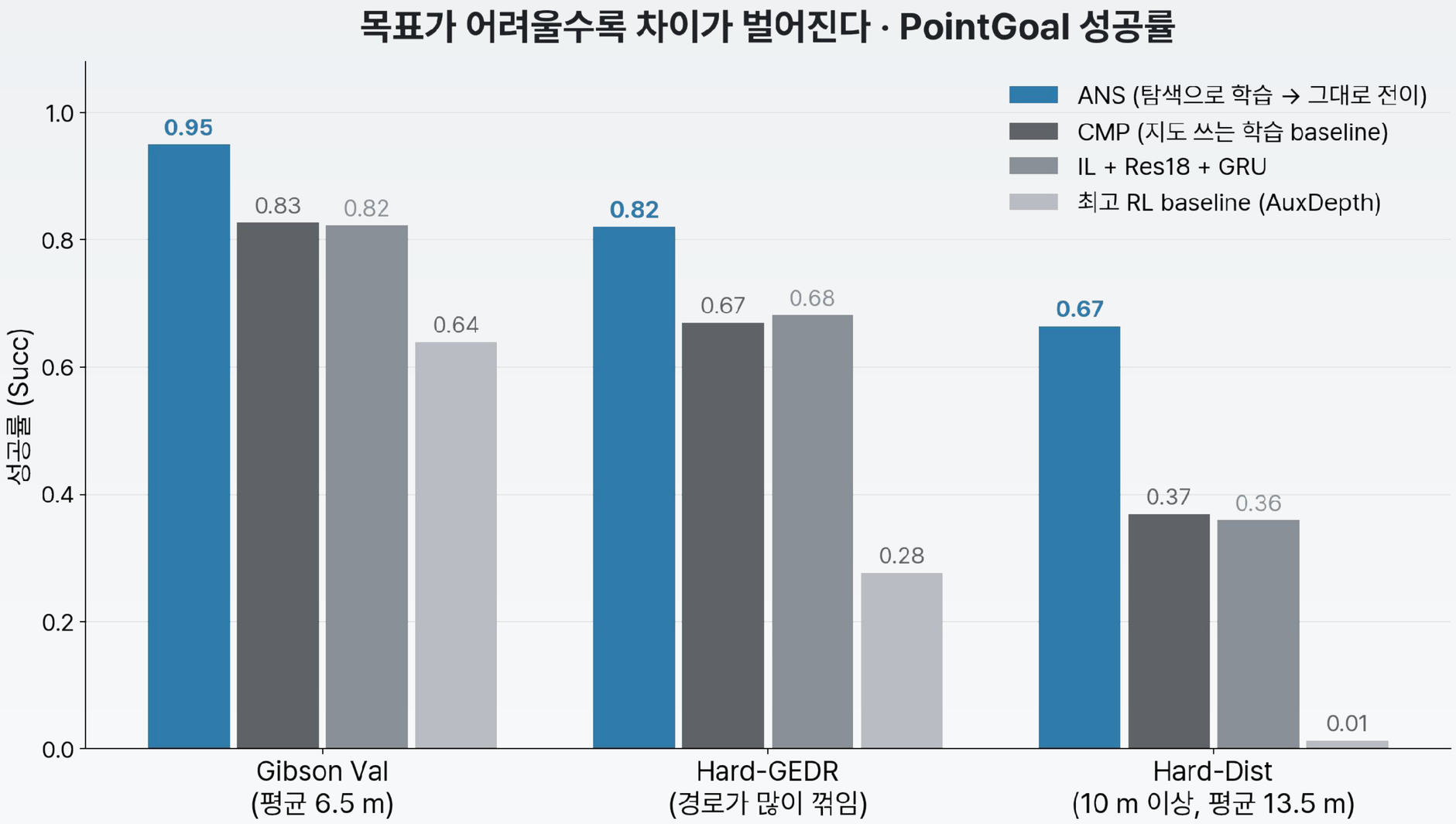
RGB-D Leaderboard

Rank	Team	SPL
1	Arnold	0.948
2	Pansy	0.927
3	titardrew	0.868
4	Hiccup	0.846
10	hela-ppo-baseline	0.569

Chaplot et al., ICLR 2020, Fig 6

By Goal Difficulty

The harder the goal, the wider the gap



Data: Chaplot et al., ICLR 2020, Table 3

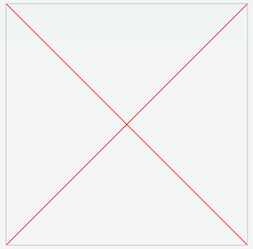
05

Limitations



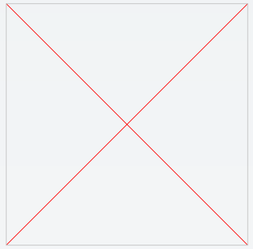
Limitations

Three limitations



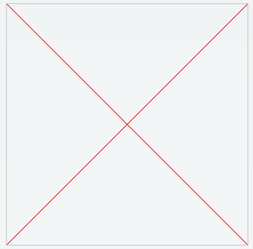
Drift accumulates

No loop closure



Fixed map size

Train 24 m · Eval 48 m



Sim ground truth

Sim-to-real gap

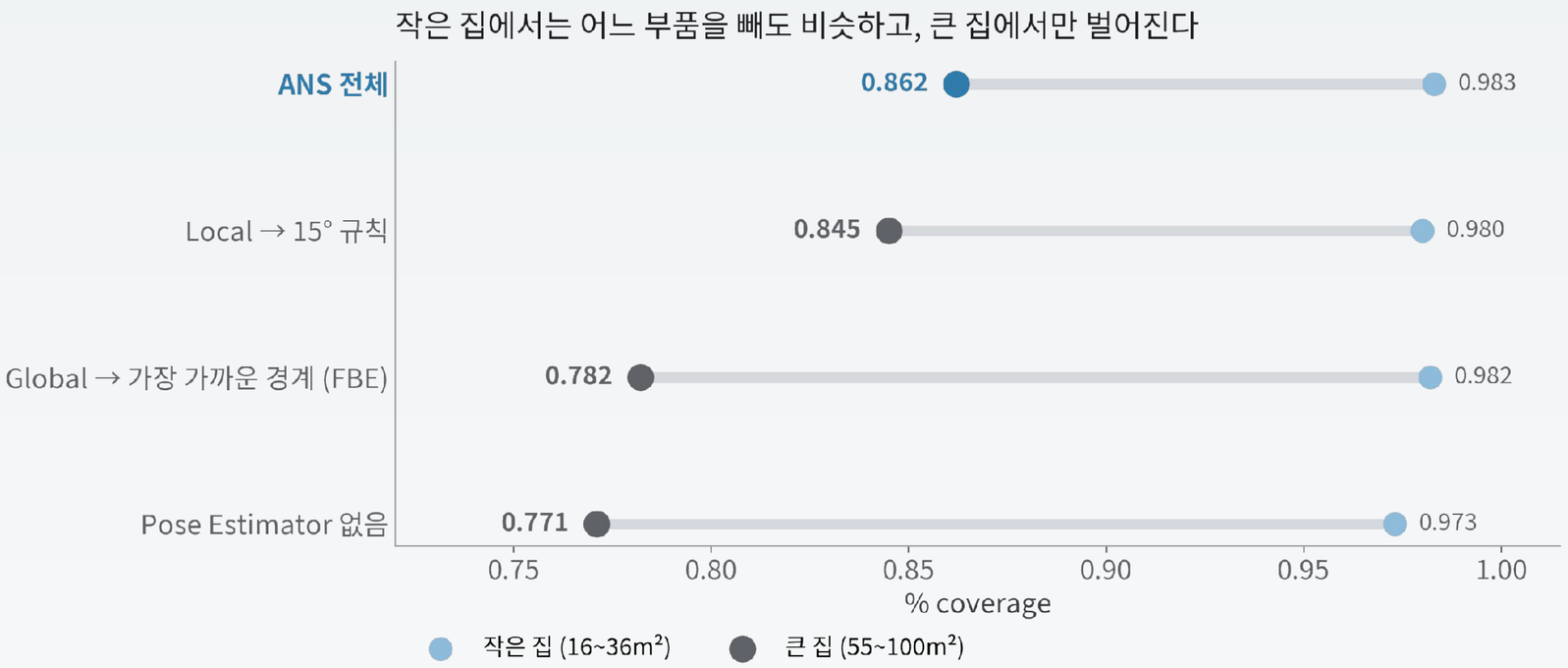
Q&A

Questions



Backup B1

Does ANS win only because of GT pose?



0.916

ANS w/o Pose Estimator

0.789

Best baseline

Chaplot et al., ICLR 2020

Backup B3

Why PPO?

Cheap, parallel data

Simulator · 72 houses at once

Short horizon

40 decisions / episode

Continuous output

Goal (x, y) on the map

Fair comparison

Same PPO · same reward as baselines