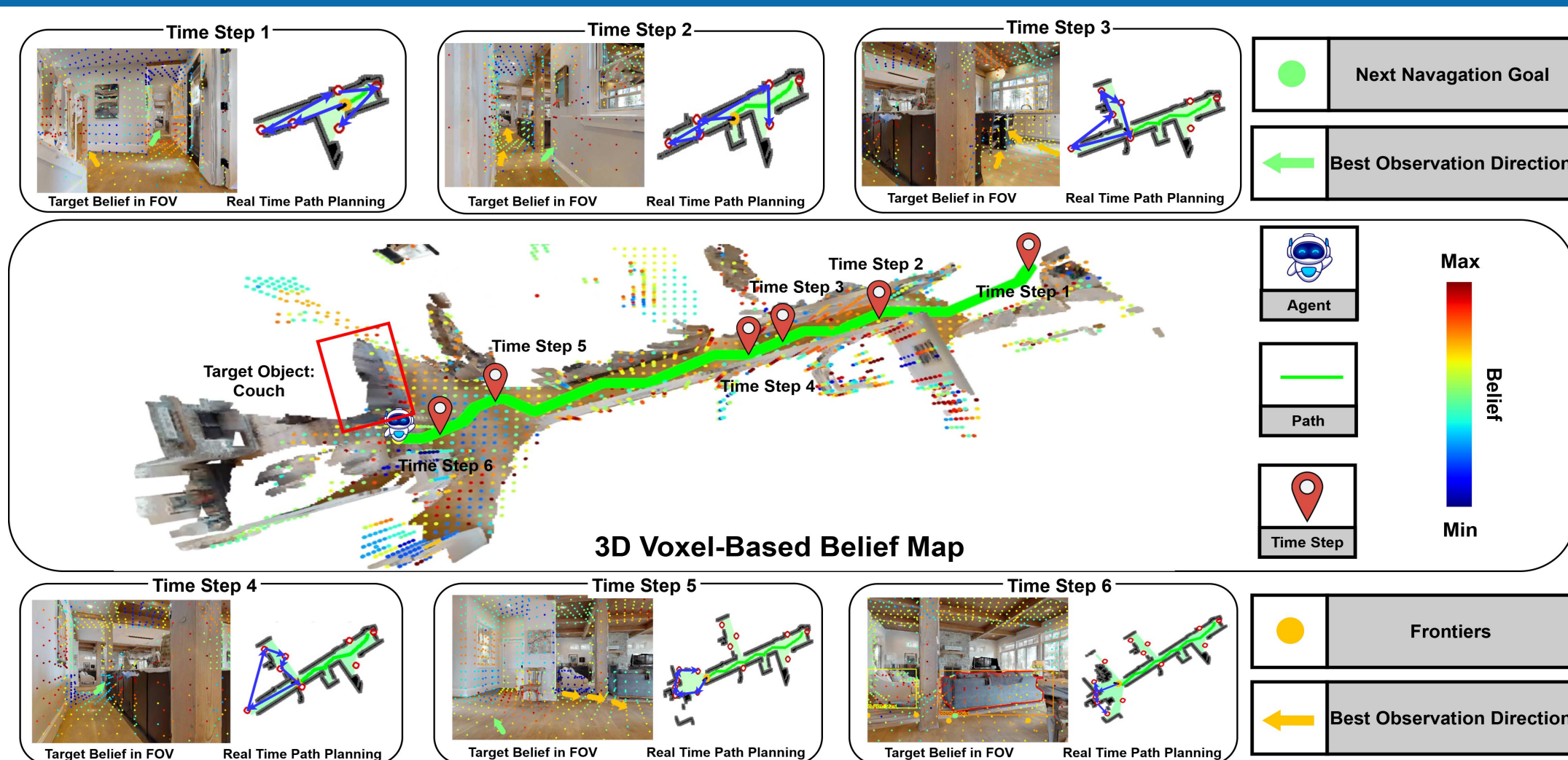


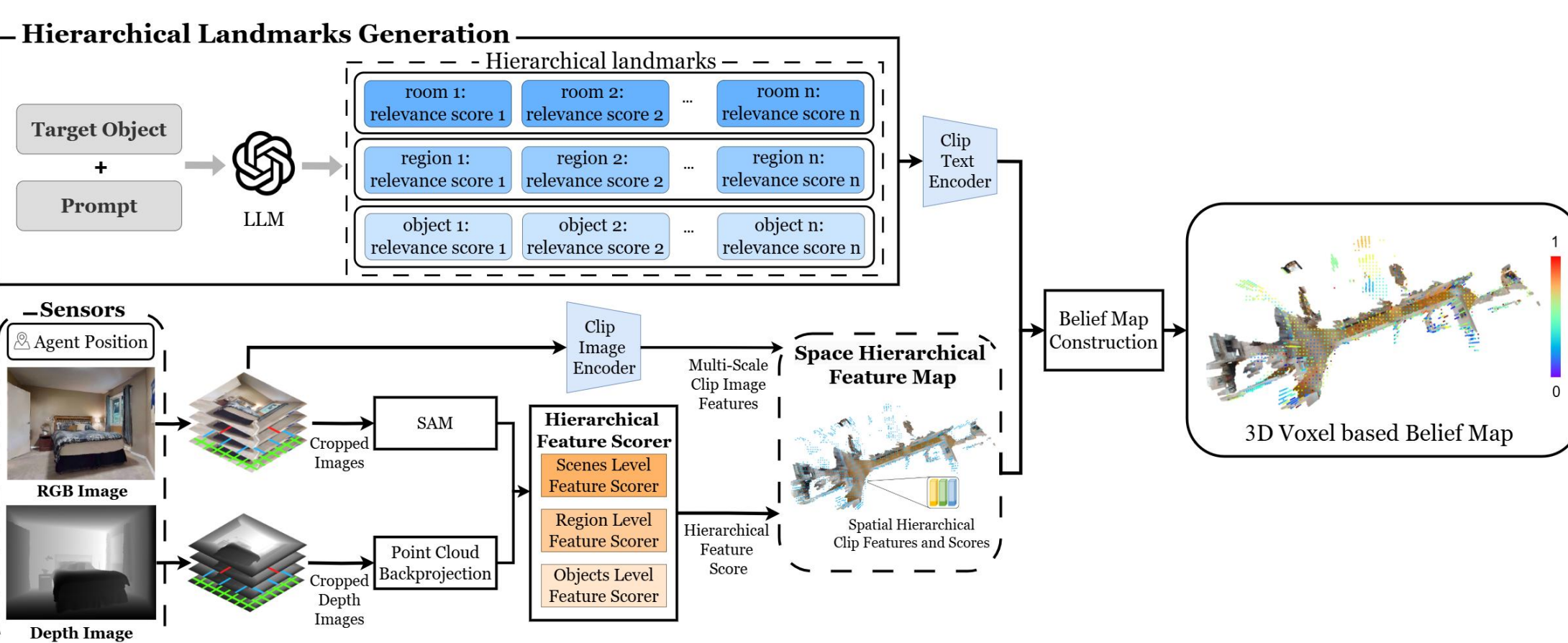


Introduction



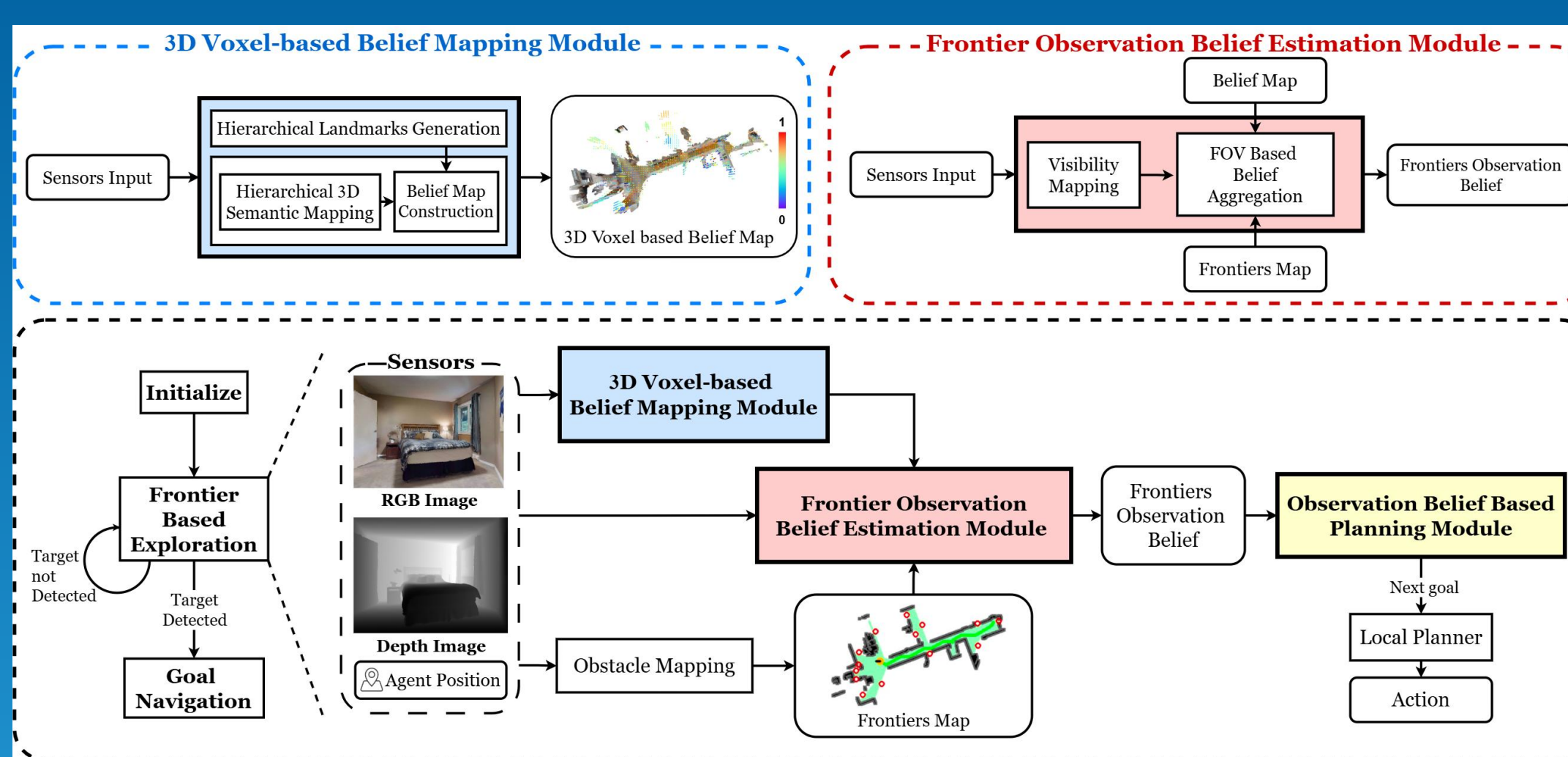
- ◆ Conventional object-goal navigation requires heavy task-specific training, BeliefMapNav builds an **online 3D voxel belief map** fusing hierarchical visual semantics, visibility cues, and LLM commonsense, and plans by minimizing expected search distance—enabling **zero-shot**, efficient search with superior cross-scene generalization.

Belief Map

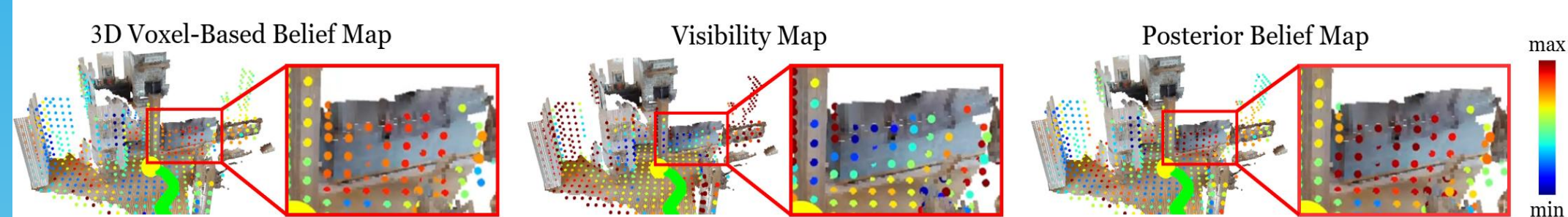


- ◆ Hierarchical Spatial Semantics: Voxelize the scene from RGB-D and assign multi-level semantics to each voxel.
- ◆ Commonsense from LLMs: Using the environment description, we prompt the LLM about the target object; it returns multi-level landmark candidates with relevance scores.
- ◆ Belief Map Generation: Fuse per-voxel semantics with LLM priors to yield a 3D prior probability of the target's spatial location.

Method



- ◆ 3D voxel belief mapping: fusing live RGB-D/SLAM with the 3D hierarchical semantic map and LLM-prompted landmarks to maintain an online belief map.
- ◆ frontier observation belief estimation: aggregating belief and visibility (FOV ray-casting) over candidate frontiers to score their expected observations.
- ◆ observation-belief planning: selecting the next goal that minimizes expected search distance and issuing navigation actions. The loop repeats until the target is detected, then the agent drives to the object.



- ◆ Visualization of the prior belief map, visibility map, and the posterior belief map, with an enlarged section highlighting the target object. Red color represents a high level of belief in the existence of the target.

$$\pi^* = \operatorname{argmin}_{\pi \in S_n} \sum_{i=1}^n \left(\sum_{k=1}^i d_{A^*}(f_{\pi_{k-1}}, f_{\pi_k}) \right) P_{\text{obs}}(f_{\pi_i})$$

- ◆ The belief-based planning objective, which ensures efficient search and reduces unnecessary exploration.

Experiments

- ◆ Our BeliefMapNav achieves leading performance across multiple datasets and metrics, outperforming most prior methods.

Method	Unsupervised	Zero-shot	HM3D		MP3D		HSSD	
			SR↑	SPL↑	SR↑	SPL↑	SR↑	SPL↑
Habitat-Web [41]	×	×	41.5	16.0	31.6	8.5	-	-
OVRL [42]	×	×	-	-	28.6	7.4	-	-
ProcTHOR [43]	×	×	54.4	31.8	-	-	-	-
SGM [44]	×	×	60.2	30.8	37.7	14.7	-	-
ZSON [24]	×	✓	25.5	12.6	15.3	4.8	-	-
PSL [45]	×	✓	42.4	19.2	18.9	6.4	-	-
PixNav [11]	×	✓	37.9	20.5	-	-	-	-
VLFM [7]	✓	✓	52.5	30.4	36.4	17.5	-	-
ESC [3]	✓	✓	39.2	22.3	28.7	14.2	38.1	22.2
Cows [13]	✓	✓	-	-	9.2	4.9	-	-
L3MVN [14]	✓	✓	50.4	23.1	34.9	14.5	41.2	22.5
ImagineNav [46]	✓	✓	53.0	23.8	-	-	51.0	24.9
VoroNav [28]	✓	✓	42.0	26.0	-	-	41.0	23.2
GAMap [8]	✓	✓	53.1	26.0	-	-	-	-
OpenFMNav [47]	✓	✓	52.5	24.1	37.2	15.7	-	-
SG-Nav [12]	✓	✓	54.0	24.9	40.2	16.0	-	-
UniGoal [48]	✓	✓	54.5	25.1	41.0	16.4	-	-
InstructNav [9]	✓	✓	58.0	20.9	-	-	-	-
BeliefMapNav	✓	✓	61.4	30.6	37.3	17.6	65.2	32.1

- ◆ Visualization of Navigation Process:

