NEURAL INFORMATION PROCESSING SYSTEMS

Active Test-time Vision-Language Navigation

Heeju Ko · Sung June Kim · Gyeongrok Oh · Jeongyoon Yoon · Honglak Lee · Sujin Jang · Seungryong Kim · Sangpil Kim*

KOREA UNIVERSITY · UNIVERSITY OF MICHIGAN · SAMSUNG · KAIST

Introduction

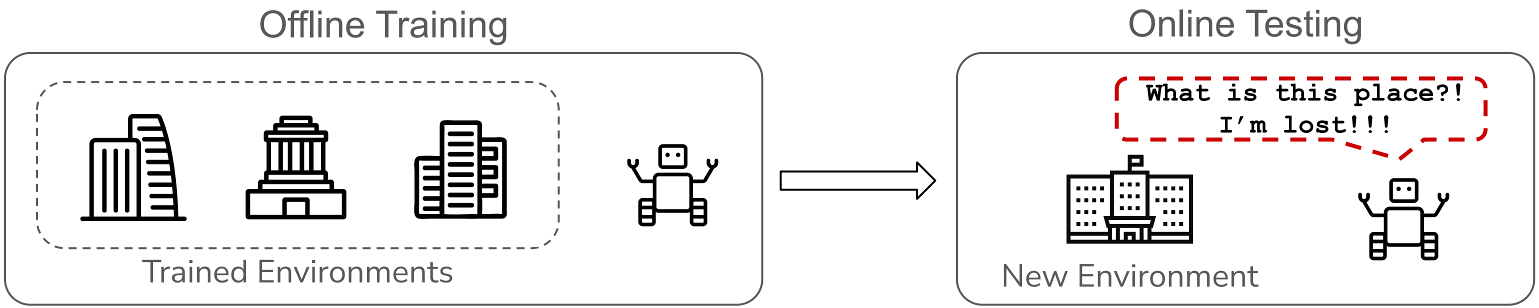
Vision-Language Navigation

- Vision-Language Navigation (VLN) is a fundamental task of connecting **human interactions** with **robotic AI systems**
- **Multimodal task** of understanding natural language instruction to navigate visual environment.

Introduction

Problem formulation

- The discrepancy between **offline training and online testing** environment hinders test-time performance.



- Possible Solutions

① Enhance generalization during training

→ Limited to 'anticipating' domain shifts. Can not cover all real-world diversity.

② Zero-shot navigation using LLMs

→ Low performance. Requires fine-tuning on navigation data for reliable performance.

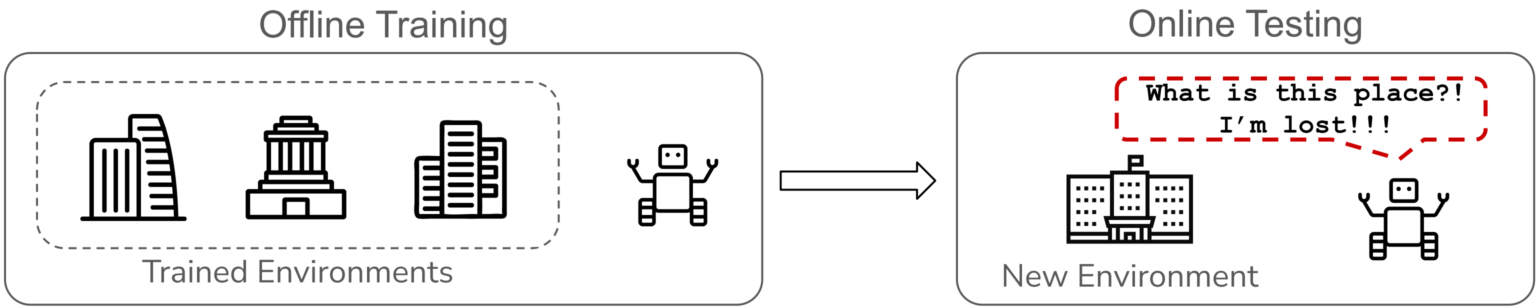
③ Test-time adaptation using entropy minimization

→ Blindly minimizing entropy without **correct signal** causes overconfidence & error accumulation

Introduction

Problem formulation

- The discrepancy between **offline training and online testing** environment hinders test-time performance.



- Possible Solutions

① Enhance generalization during training

→ Limited to 'anticipating' domain shifts. Can not cover all real-world diversity.

② Zero-shot navigation using LLMs

Can't navigation agents actively interact with humans at test time?

③ Test-time adaptation using entropy minimization

→ Blindly minimizing entropy without **correct signal** causes overconfidence & error accumulation

Introduction

Research objective

In this work, we aim to explore how **active learning** can benefit test-time navigation.

Specifically, we study…

- how active learning can be utilized in the context of test-time vision-language navigation
- how we can improve entropy as a reliable test-time signal
- how agent can learn in the absence of human interaction

Method

Active Test-time Navigation

Ⓘ Preliminary

Active Learning → A learning algorithm that interactively query a human user (or some other information source), to label new data points

Example.

- Traditional AL uses uncertainty sampling with metrics like entropy, margin, and least confidence.
- Initially applied to simple classification, later extended to complex real-world tasks.
- Recently expanded into Test-Time Adaptation (TTA) — models adapt during inference using uncertainty estimates.

Active Test-time Adaptation → incorporating limited labeled test instances to enhance overall test time performances.

Method

Active Test-time Navigation

② What is different?

Active Test-time Navigation requires several constraints:

- Latency: 'Labeling' in navigation task implies a step-wise expert demonstration. Even if it's active learning, labeling the correct action for each time-step is infeasible in test time.

- Accessibility: The interface for human input should be intuitive and require minimal expertise or effort from the end user.

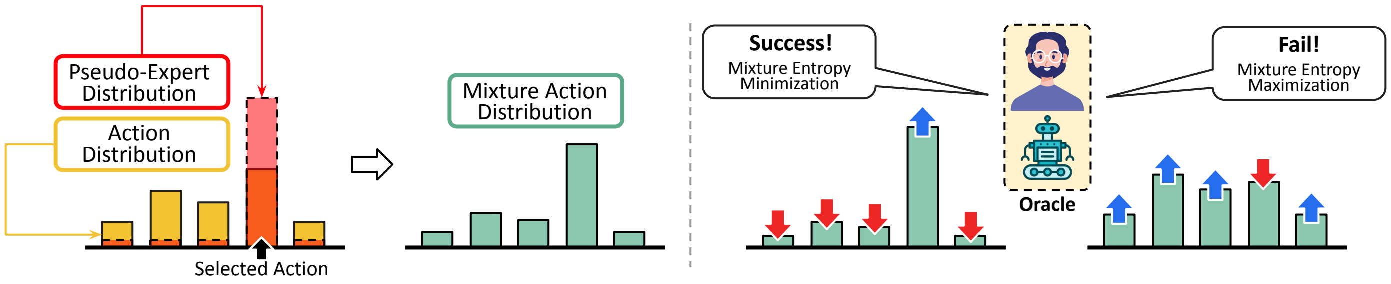
Therefore, in **Active Test-time Navigation**, the oracle provides the agent with…

"Episodic Evaluation"

Method

Active Test-time Navigation

② How do we use it?



(a) Mixture Action Distribution (b) Mixture Entropy Optimization

Mixture Action Distribution

- combine predicted action distribution with pseudo-expert distribution via convex combination, enhancing the clarity (sharpness) of actions.

Mixture Entropy Optimization

- minimizing for a successful navigation and maximizing for a failed navigation can help the agent calibrate its confidence more effectively, **reinforcing correct behaviors** while **discouraging erroneous ones**.

Method

Mixture Entropy Optimization

Mixture Action Distribution Formulation

Mixture action distribution

$$L_{ent} = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \Rightarrow \boxed{L_{ent}^{*} = -\sum q_{mix}(x_i) \log q_{mix}(x_i)} \qquad \boxed{q_{mix}(x_i) = \lambda q_{pseudo}(x_i) + (1-\lambda)\pi_\theta(x_i)}$$

Original entropy term

Mixture entropy term

*pseudo-expert dist.

predicted prob dist.

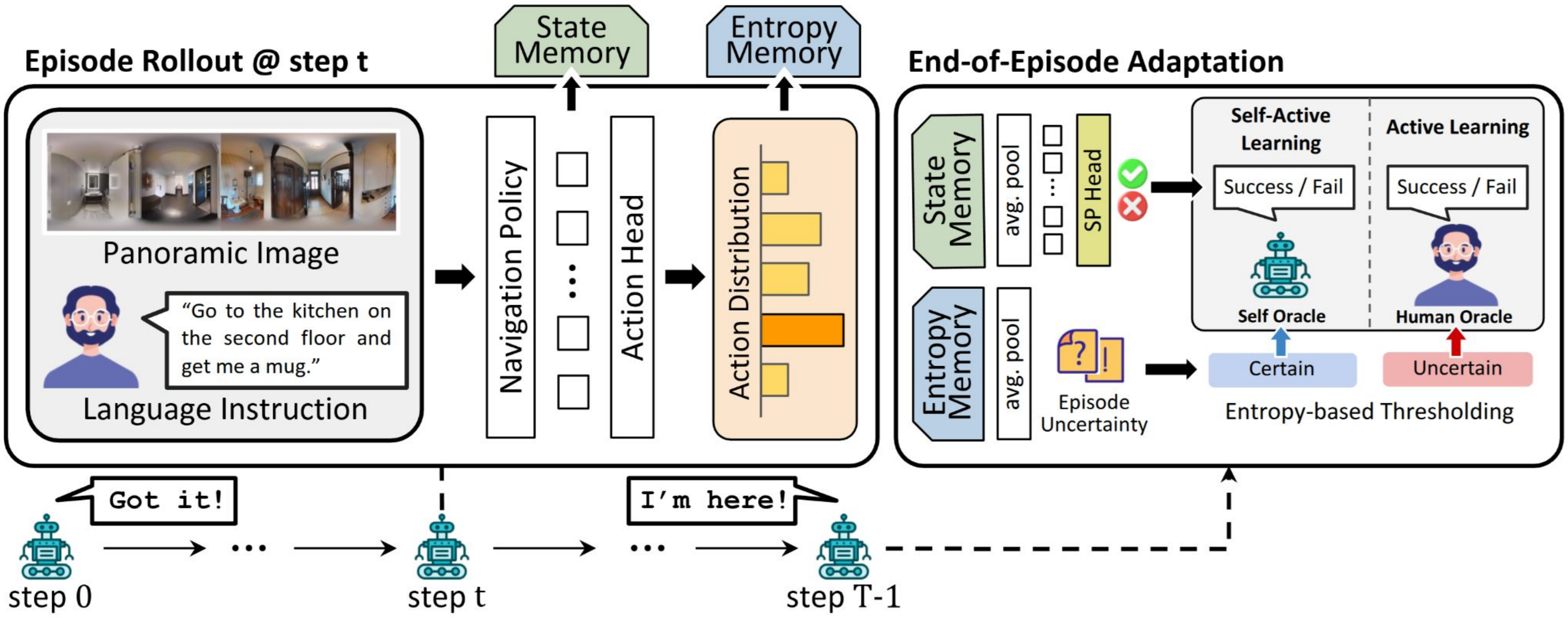
- **Pseudo-Expert Distribution.** treat the action taken by the agent as an expert demonstration
- **Mixture Action Distribution.** combine the original predicted probability distribution and the pseudo-expert distribution with convex combination.
- **Mixture Entropy Optimization.** the adaptation optimizes the sharpened distribution, strongly reinforcing successful predictions while explicitly penalizing failed predictions, thus flattening the distribution.

Method

Self Active Learning

- Injecting a self-prediction head to predict whether or not the navigation was a success or not.

- Train through test-time stream and use it when human query is inactive.



Experiments

Results on Vision-Language Navigation Datasets

REVERIE Dataset

Methods	Val Seen				Val Unseen				Test Unseen			
	OSR ↑	SR ↑	SPL ↑	RGSPL ↑	OSR ↑	SR ↑	SPL ↑	RGSPL ↑	OSR ↑	SR ↑	SPL ↑	RGSPL ↑
HAMT [5]	47.65	43.29	40.19	25.18	36.84	32.95	30.20	17.28	33.41	30.40	26.67	13.08
w/ TENT† [7]	46.03	43.43	40.78	25.81	32.60	30.56	28.23	14.48	25.06	23.73	21.78	10.82
w/ FSTTA† [37]	48.21	42.87	39.56	24.58	36.78	32.89	30.51	17.20	33.39	30.39	26.65	13.61
w/ ATENA (**Ours**)	52.92	57.34	48.08	29.60	38.85	34.00	30.96	17.51	38.19	32.55	28.38	14.32
DUET [4]	73.86	71.75	63.94	51.14	51.07	46.98	33.73	23.03	56.91	52.51	36.06	22.06
w/ TENT	73.72	71.89	64.06	50.41	51.43	47.55	33.99	23.32	57.12	52.61	36.17	22.16
w/ FSTTA	75.59	75.48	65.84	52.23	56.26	54.15	36.41	23.56	58.44	53.40	36.43	22.40
w/ ATENA (**Ours**)	85.52	84.33	74.31	59.99	71.88	68.11	45.82	31.26	57.74	54.28	40.70	25.01
GOAT† [60]	82.36	80.74	73.44	58.82	57.97	53.82	37.52	27.00	61.44	57.72	40.53	26.70
w/ TENT†	82.43	80.74	73.47	58.75	57.68	53.51	37.49	26.99	62.00	57.28	39.82	26.97
w/ FSTTA†	82.36	80.74	73.42	58.82	57.94	53.79	37.50	26.95	62.35	57.52	39.49	26.82
w/ ATENA (**Ours**)	85.03	83.35	76.45	61.60	70.29	67.66	53.15	39.80	64.26	62.03	46.82	31.54

R2R Dataset

Methods	Val Seen				Val Unseen			
	TL ↓	NE ↓	SR ↑	SPL ↑	TL ↓	NE ↓	SR ↑	SPL ↑
DUET [4]	12.33	2.28	79	73	13.94	3.31	72	60
w/ FSTTA [37]	13.39	2.25	79	73	14.64	3.03	75	62
w/ ATENA (**Ours**)	11.27	2.18	80	75	12.31	2.90	75	66
BEVBert [59]	13.56	2.17	81	74	14.55	2.81	75	64
w/ FSTTA†	12.28	2.31	80	75	13.96	2.89	74	63
w/ ATENA (**Ours**)	10.79	2.26	82	78	12.22	2.78	76	68
GOAT† [60]	11.87	1.70	84.52	79.60	13.43	2.33	77.91	67.34
w/ FSTTA†	11.67	1.65	84.92	80.08	13.26	2.32	77.99	67.48
w/ ATENA (**Ours**)	11.66	1.64	85.01	80.13	12.52	2.27	79.01	69.30

R2R-CE Dataset

Methods	Val Seen					Val Unseen				
	TL ↓	NE ↓	OSR ↑	SR ↑	SPL ↑	TL ↓	NE ↓	OSR ↑	SR ↑	SPL ↑
ETPNav [6]	11.78	3.95	72	66	59	11.99	4.71	65	57	49
w/ FSTTA† [37]	11.35	3.93	72	66	59	11.57	4.77	64	57	49
w/ ATENA (**Ours**)	10.81	3.86	72	67	61	12.89	4.53	66	58	49
BEVBert [59]	13.98	3.77	73	68	60	13.27	4.57	67	59	50
w/ FSTTA	14.07	4.11	74	69	60	13.11	4.39	65	60	51
w/ ATENA (**Ours**)	11.31	3.24	75	71	64	13.48	4.50	67	60	51

Conclusion

Contribution & Novelty

1. First to explore how active human-agent interaction at test time improves navigation performance.
2. Propose Mixture Entropy Optimization to explicitly reward or penalize the actions.
3. Propose Self Active Learning, where agent queries itself for labels at relatively certain navigations.

→ **Experimental results show a substantial improvement over the baseline policy.**