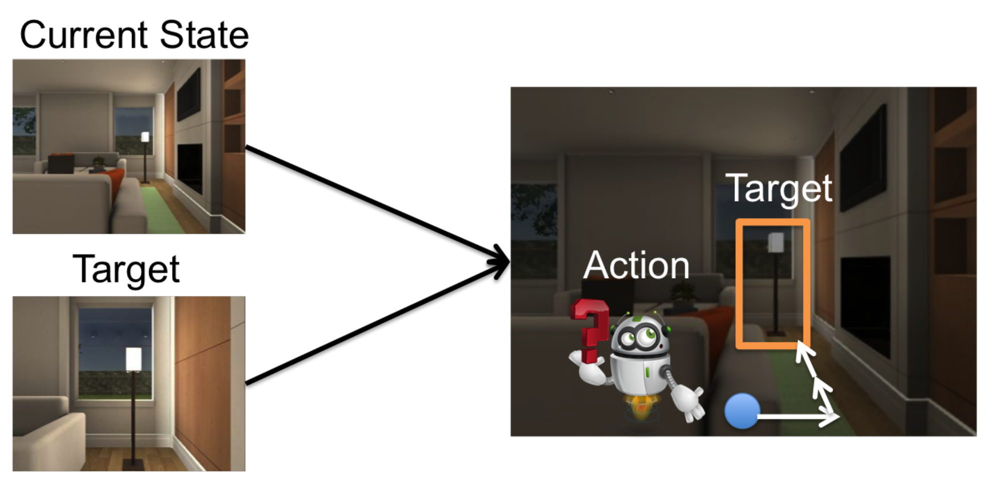
Indoor Target-driven Visual Navigation Using Imitation Learning

Yu Guo[1], Xin Zheng[1]
[1]Department of Electrical Engineering

Task

- Navigate the robot to a target image
- Input: Current state (o_t) and target image
- Output: action (a_t) to take



Dataset [1]

- No. of scenes = 20
- Avg. No. of states = 519

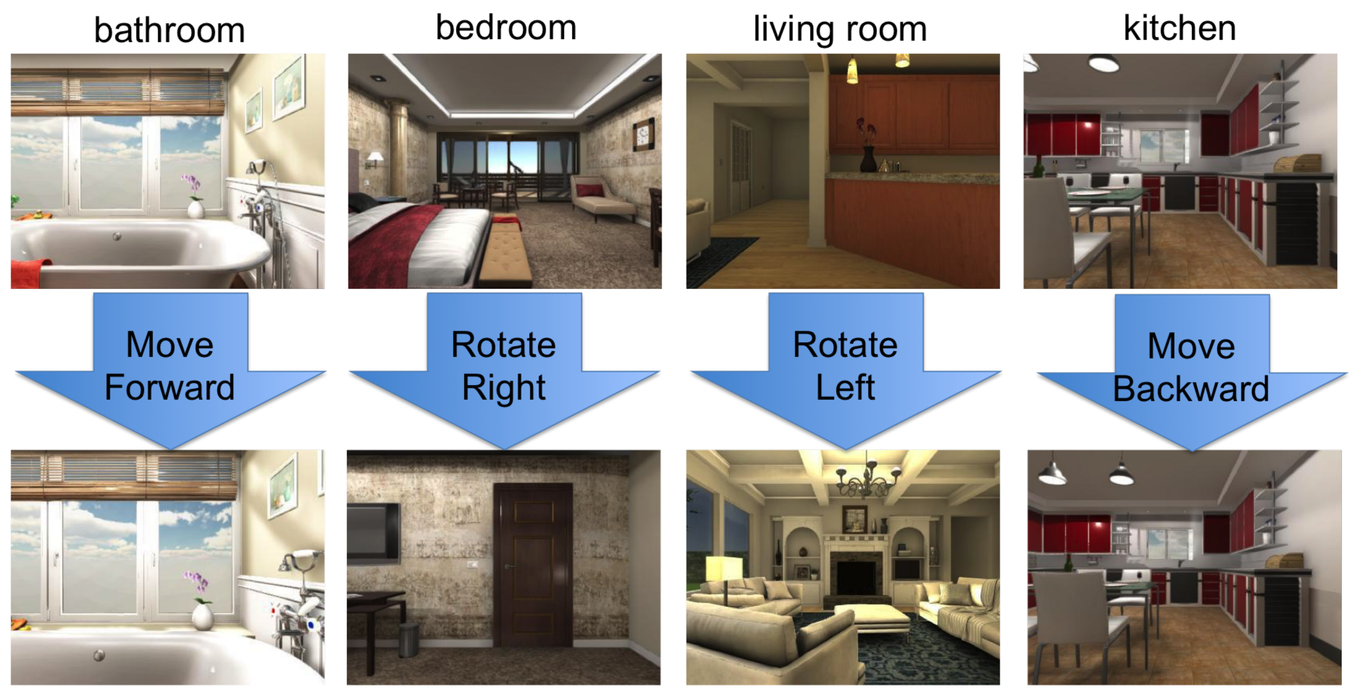
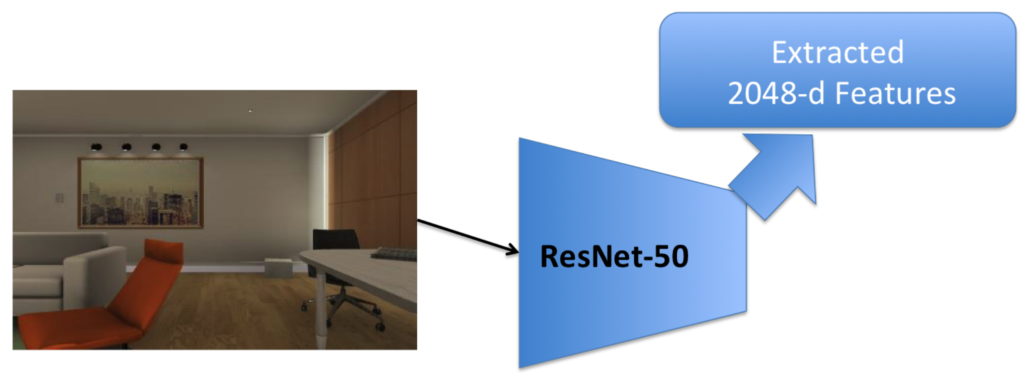
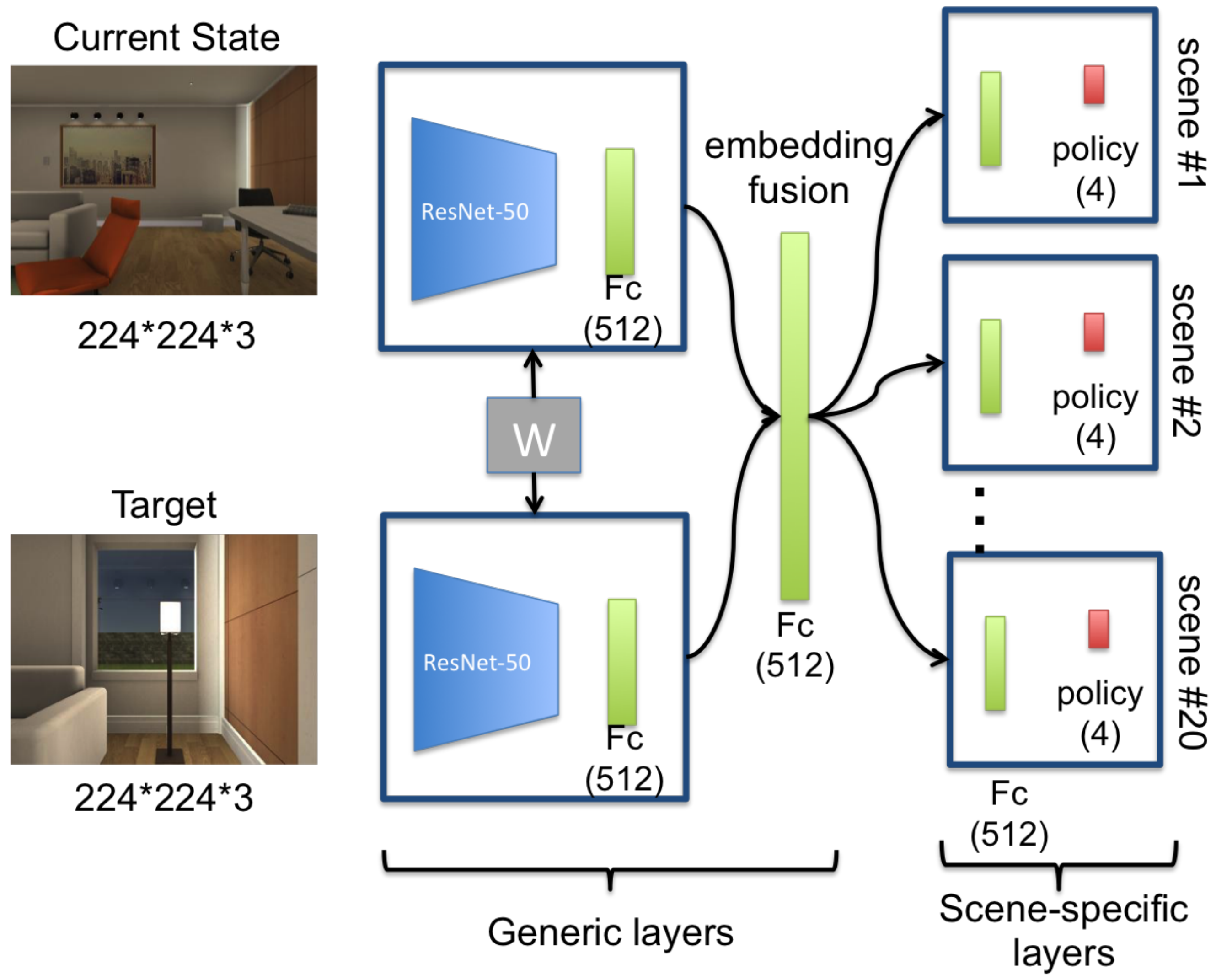


Image Preprocessing

- Extract 2048-d features through ResNet-50



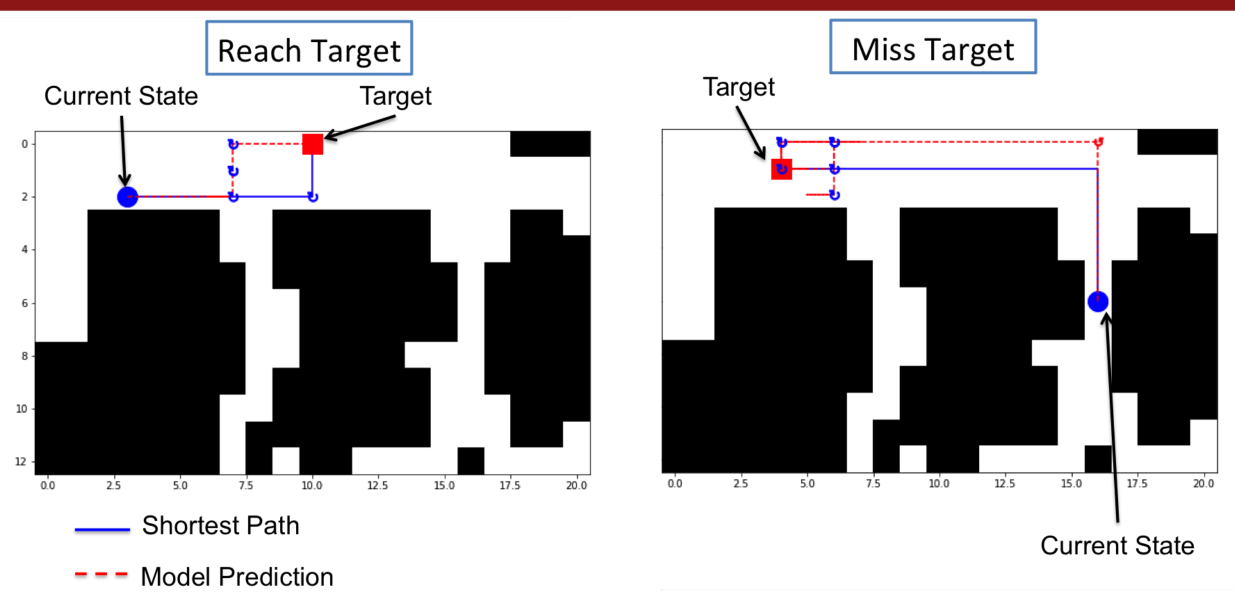
Main Architechture



Dagger (Dataset Aggregation) [2]

- train policy $\pi_\theta(a_t|o_t)$ from human data $D_{\pi^*} = (o_1, a_1, ..., o_N, a_N)$
- run policy $\pi_\theta(a_t|o_t)$ to get observations $D_\pi = (o_1, ..., o_M)$
- ask human to label dataset D_π with optimal actions a_t
- Aggregate $D_{\pi^*} \leftarrow D_{\pi^*} \cup D_\pi$
- go to first step

Qualitative Results



Quantitative Evaluation

	Scene 1	Scene 2	Scene 3	Scene 4
states	180	408	676	468
basic imitation learning				
success rate	0.93	0.80	0.96	0.90
avg. shortest length	6.58	13.04	17.60	12.78
avg. predicted length	37.57	76.76	33.92	67.68
avg. successful length	18.82	22.11	21.86	40.91
DAGGER (iteration = 1)				
success rate	0.96	0.89	0.93	0.88
avg. shortest length	6.47	12.90	17.67	14.10
avg. predicted length	19.94	45.34	41.23	51.04
avg. successful length	9.28	14.93	21.75	18.16
DAGGER (iteration = 2)				
success rate	0.98	0.98	0.92	0.92
avg. shortest length	6.76	12.44	17.71	13.59
avg. predicted length	15.29	20.98	41.56	37.38
avg. successful length	9.48	15.29	20.10	15.58

*all paths cut off at 300 steps

Conclusion

- The robot follows shortest path through imitation learning
- DAGGER generally improves the performance, stabilize the behavior of the robot

Future Work

- Parallelize the DAGGER module to accelerate data collection
- Generalize to new tasks and new scenes
- Investigate other model structures such as CNN and deep reinforcement learning

References

[1] Yuke Zhu, Roozbeh Mottaghi, Eric Kolve, Joseph J Lim, Abhinav Gupta, Li Fei-Fei, and Ali Farhadi. Target-driven visual navigation in indoor scenes using deep reinforcement learning. *arXiv preprint arXiv:1609.05143*, 2016.

[2] Stéphane Ross, Geoffrey J Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *AISTATS*, volume 1, page 6, 2011.