



새로운 강화학습 환경에서의 부분 관측 마르코프 결정 프로세스 학습

양현서°, 장병탁 (Biointelligence Lab)



I. Contributions

- 마인크래프트 기반 새로운 강화학습 환경 CraftGround
- 최신 버전 지원 및 업데이트 용이
- 고성능 (300 FPS) 지원: protobuf와 unix domain socket 사용
- 다양한 관측 공간 제공: 소리, 통계, 사망 원인 등
- POMDP 문제 해결을 위한 효과적 환경 제공

II. 새로운 강화학습 환경 CraftGround

	Ours	Malmö
Version	1.19.4	1.11.2
Speed	300TPS	~120TPS
Obs/Vision	O	O
Obs/Sound	O	X
Obs/Binocular	O	X

i. Observation Space

이미지, 좌표, 시선 방향, 체력, 배고픔, 사망 여부, 소지 아이템, 시선 블록 및 엔티티, 소리 자막, 상태 효과, 통계, 가시 엔티티, 주변 엔티티, 낚싯대 상태, 경험치, 시각, 사망 메시지, 양안시 시각

ii. Action Space

- MineDojo[1] 의 행동 공간과 같음 (Crafting 미지원)

III. 실험

i. 환경

- 무한 평지에서 시작하며 적대적 몬스터인 허스크가 공격
- 목표: 오랜 시간 생존
- 비전 입력: 114x64, HUD 가림
- 소리 입력: (dim × (소리의 경우의 수) + 3) 길이의 벡터
 - ▶ 음원 상대 좌표
 - ▶ 에이전트 yaw의 cos, sin [2]
 - ▶ 피격 여부
- 매 스텝마다 1/50 확률로 15-30 거리에서 10마리 허스크 소환
- 행동 공간: [정지, 전진, 후진, 좌우 이동, 좌우 회전]
- 에피소드 종료 조건: 400틱 도달 또는 사망

ii. 보상

- 피해 시 -0.1
- 피해 없을 시 매 틱마다 0.5
- 사망 시 -1
- 1블록 이내 허스크 수 H_1 , 2블록 이내 H_2 , 5블록 이내 H_5 일 때

$$-0.01(H_1 + 0.5H_2 + 0.25H_5) \quad (1)$$

iii. 모델 및 실험

	Multimodal	Recurrent
kcc-husk-vision-dqn	X	X
kcc-husk-bimodal-dqn-final	O	X
kcc-husk-vision-drqn-final	X	O
kcc-husk-bimodal-drqn-final	O	O

- 모델: Dueling DQN [3]
- 학습 에피소드: 3000
- 평가: 10회 에피소드마다
- 평가 지표: 에피소드 길이, 획득 보상

IV. Result & Analysis

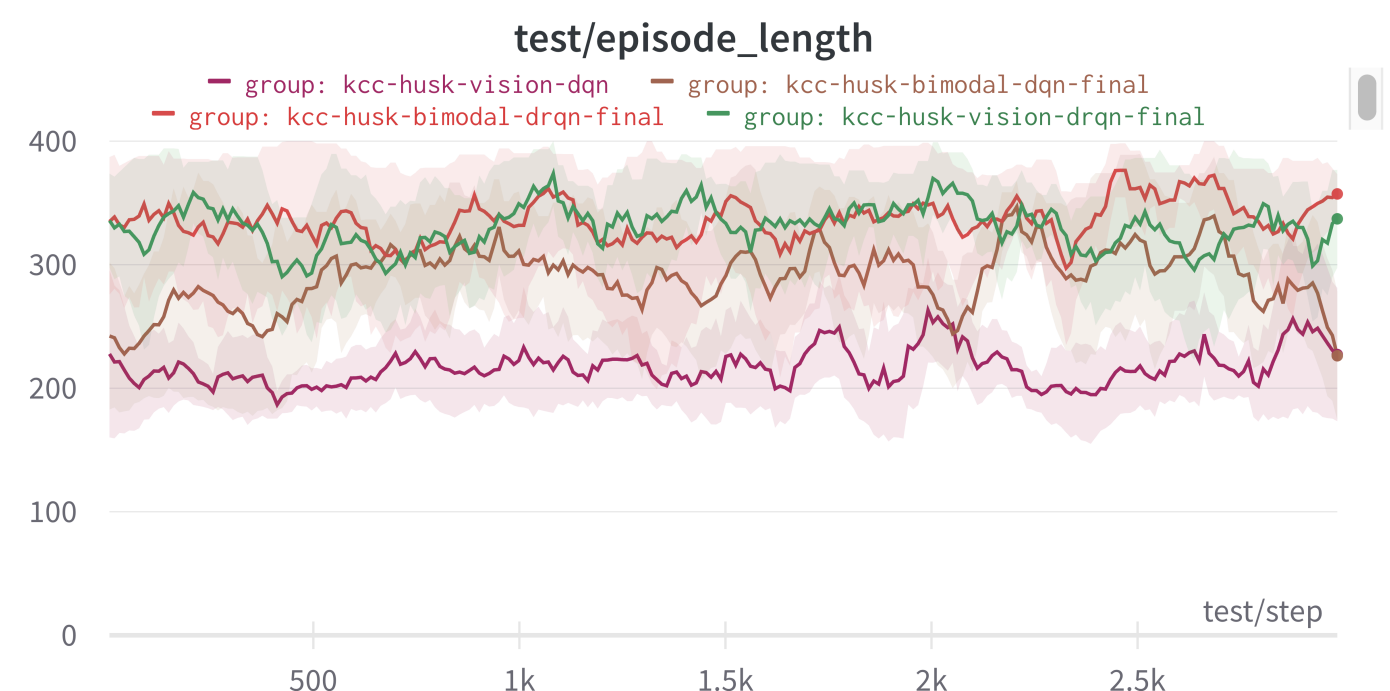


Figure 1: 에피소드별 에이전트의 최대 생존 시간을 이동평균한 그래프

- Recurrent Layer를 사용한 경우 Vision만 사용한 경우와 Bimodal 입력을 사용한 경우가 비슷한 성능을 보임
- Recurrent Layer를 사용하지 않은 경우 Recurrent Layer를 사용한 경우보다 성능이 떨어짐
 - ▶ Bimodal 입력을 사용한 경우가 Vision만 사용한 경우보다 더 나은 성능을 보임

Bibliography

- [1] L. Fan et al., "MineDojo: Building Open-Ended Embodied Agents with Internet-Scale Knowledge," CoRR, 2022, [Online]. Available: <https://arxiv.org/abs/2208.xxxxx>
- [2] Y. Zhou, C. Barnes, J. Lu, J. Yang, and H. Li, "On the Continuity of Rotation Representations in Neural Networks," CoRR, 2018, [Online]. Available: <http://arxiv.org/abs/1812.07035>

- [3] Z. Wang, N. Freitas, and M. Lanctot, "Dueling Network Architectures for Deep Reinforcement Learning," CoRR, 2015, [Online]. Available: <http://arxiv.org/abs/1511.06581>