자율 이동 로봇의 충돌 회피 모델 평가를 위한 적대적 공격 사용

김정은, 이욱, 석준희*
고려대학교, 고려대학교, *고려대학교

mine1372@korea.ac.kr, leewook94@korea.ac.kr, *jseok14@korea.ac.kr

Using Adversarial Attacks to Evaluate Collision Avoidance Models in Autonomous Mobile Robots.

Kim Jeong Eun, Lee Wook, Seok Jun Hee*
Korea Univ., Korea Univ., *Korea Univ

요 약

많은 산업에서 자동화가 점차 보편화됨에 따라 자율 이동 로봇은 다양한 분야에서 활용되고 있다. 현실에서 작업 효율성을 극대화하기 위해서 자율 이동 로봇은 짧은 시간 내에 안전하게 목적지까지 도달해야 한다. 이러한 목적을 달성하기 위해 최근 연구들은 데이터 혹은 시뮬레이터로 충돌 회피 모델을 최적화하는 접근법을 많이 사용한다. 본 논문은 학습 기반 충돌 회피 모델의 안전성을 보장하기 위해 적대적 공격을 사용하는 평가 방법을 제시한다.

Ⅰ. 서 론

자율 이동 로봇(Autonomous Mobile Robot)은 사람이 수동으로 하는 작업들을 자동화하여 생산성을 향상시키고 작업 안전성을 높일 수 있다. 예를 들어 제조 분야에서는 생산 라인의 자동화 비율을 높여 생산 속도를 일정하게 유지할 수 있고, 물류 분야에서는 무거운 물건을 운반할 수 있어 부상 위험이 낮아진다. 기존에 많이 사용되던 무인 운반 차량(Automated Guided Vehicle)과의 가장 큰 차이점은 정해진 길을 따라가는 것이 아닌 자율적으로 길을 찾는 것이다. 자율성이 목적지로 가는 시간을 단축시킬 수 있지만 위험성을 증가시키기도 한다. 시간 효율성과 안전성을 모두 높이기 위해 최근 연구들은 데이터 혹은 시뮬레이터를 사용하여 충돌 회피 모델을 최적화하는 접근법을 사용하였다. 이러한 학습 기반 모델의 치명적인 단점은 모든 상황에 대해 안전성을 보장할 수 없다는 것이다. 기존 충돌 회피 모델의 평가 방식은 적은 수의 상황에서 실험하거나 랜덤한 상황 속에서 평가를 진행한다. 이와 같은 방식은 실패 사례를 발견하기 어렵다. 본 논문은 학습 기반 충돌 회피 모델의 안전성을 보장하기 위해 적대적 공격을 사용한 평가 방법을 제시한다.

Ⅱ. 학습 기반 충돌 회피

학습 기반 충돌 회피 모델은 동적 환경에서 자율 이동 로봇이 장애물을 피하면서 빠르게 목적지까지 도달할 수 있도록 널리 연구되고 있다. 모델의 학습을 위해 사용되는 방식은 지도 학습, 모방 학습, 강화 학습[1][2] 등 다양한 접근법이 존재한다. 학습 기반 모델은 대부분의 상황에서 기존 방식보다 높은 시간 효율성과 안전성을 보여주지만 모든 상황에 대하여 이를 보장한다고 보기 어렵다.

학습 기반 모델의 견고함과 일반화는 해당 모델이 실패하는 경우인 실패 사례를 찾음으로써 개선될 수 있다. 이러한 예외를 찾기 위해 충돌 회피 모델은 고정된 상태에서 적대적 공격 기반 평가가 사용될 수 있다.

Ⅲ. 적대적 공격 기반 평가

적대적 공격을 사용하는 가장 대표적인 연구는 적대적 생성 신경망(Generative Adversarial Networks)이 있다. 분류 문제에서 적대적 공격이 의미하는 바는 모델이 잘못된 분류를 만드는 방식으로 기존 데이터를 수정하는 것이다. 적대적 공격 기반 평가는 검증 대상이 되는 모델을 고정한 채로 실패를 계속하여 유도하는 것이다. 기존 평가 방법은 대표적인 상황을 구성하고 실험한다. 하지만 이러한 방식은 너무 적은 케이스만 테스트하기 때문에 신뢰성이 낮다. 많은 상황에서 검증 모델을 테스트하기 위한 방법은 랜덤하게 다양한 상황을 구성하는 것이지만 시간이 오래 걸린다는 단점이 있다.

검증 모델의 실패 사례를 찾도록 적대적 공격을 수행하는 에이전트가 심층강화학습으로 학습되면서 학습 기반 모델의 강건성과 일반성을 면밀히 평가하는 방식이 존재한다[3]. 에이전트를 최적화하는 수식은 (1)과 같다. 검증 대상 모델인 $f_{\theta}(\cdot)$의 모델 파라미터 θ를 고정하고 손실함수인 $\mathcal{L}$을 최대로 만드는 것이 적대적 공격을 수행하는 에이전트의 목적함수이다. 에이전트가

목적함수를 최대화하는 과정에서 기존 데이터 x_i에 ϵ보다 크기가 작은 노이즈를 더하여 얻어지는 실패 사례 $\tilde{x}$를 찾아낸다. 이 방식을 도입한 자율 이동 로봇의 충돌 회피 모델 평가는 기존 평가 방식보다 실패 사례를 잘 찾아낼 것으로 예상된다.

$$min_\theta \frac{1}{n}\sum_{i=1}^{n} \max_{\tilde{x}:||x_i-\tilde{x}||<\epsilon} \mathcal{L}(y_i, f_\theta(\tilde{x})) \quad (1)$$

Ⅵ. 실험

검증 대상이 되는 충돌 회피 모델은 LM-SARL[4]을 사용하여 학습한 뒤 모델 파라미터를 고정하였다. 적대적 공격을 하는 에이전트(즉, 동적 장애물)는 SAC[5]로 학습하였고 강화학습을 위한 시뮬레이터는 Crowd Sim[6]을 사용하였다. 에이전트의 보상 시스템은 (2)와 같다. 첫번째 항 $-r^t_{Robot}$ 은 적대적 공격을 하기 위해서 충돌 회피 모델이 받는 보상의 부호를 바꿔서 제공했다. 두번째 항 $r^t_{Close\_Goal}$ 은 동적 장애물이 로봇의 목적지에 가깝게 있을 때 패널티를 주어 무의미한 결과가 나오는 것을 방지했다. 마지막 항 $r^t_{Reaching\_Time}$ 은 로봇이 목적지에 도달하는 시간이 길어질수록 더 큰 보상을 제공했다.

$$R(s_t, a) = [-r^t_{Robot} + r^t_{Close\_Goal} + r^t_{Reaching\_Time}] \quad (2)$$

충돌 회피 모델의 강건성과 일반성의 테스트 방식은 총 3 가지를 진행하고 비교하였다.

Random : 랜덤하게 출발지와 목적지를 변경

Random + Uniform Noise : 랜덤한 출발지와 목적지에 매 스텝 노이즈를 추가

Random + Adversarial Noise : 랜덤한 출발지와 목적지에 매 스텝 적대적 공격을 통한 노이즈 추가

동적 장애물의 출발지와 목적지에 노이즈를 0.3m 이내로 추가했을 때, 평균 도착시간이 증가한 것을 Figure 1 에서 확인할 수 있다. 적대적 공격을 통한 노이즈를 학습하면 평균 도착 시간이 더욱 증가하고 충돌까지 발생시키는 것을 Table 1 에서 확인할 수 있다.

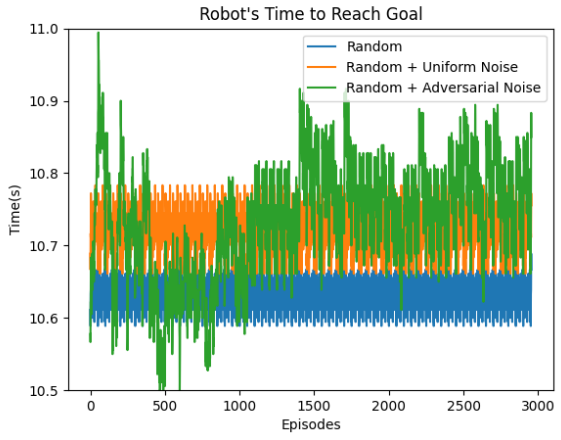

Figure 1. 검증 방식에 따른 에피소드 별 도착 시간

Methods	Success	Collision	Time(s)
Random	1.000	0.000	10.635
Random + Uniform Noise	1.000	0.000	10.720
Random + Adversarial Noise	**0.997**	**0.003**	**10.736**

Table 1. 검증 방식에 따른 충돌 비율과 평균 도착 시간

Ⅴ. 결론

자율 이동 로봇의 충돌 회피에 대한 많은 연구들은 학습 기반 접근법을 많이 사용하고 있다. 본 연구에서는 학습 기반 충돌 회피 모델의 강건성과 일반성을 검증하기 위해 적대적 공격 기반 평가를 도입하였다. 실험을 통해 적대적 공격하는 노이즈를 학습하는 것이 랜덤 테스트를 진행하는 것보다 모델의 취약성을 빠르게 발견함으로써 충돌 회피 모델을 효율적으로 평가할 수 있음을 증명하였다.

ACKNOWLEDGMENT

이 논문은 한국연구재단의 지원을 받아 수행된 연구입니다(NRF-2022R1A2C2004003).

참 고 문 헌

[1] L. Tai, J. Zhang, M. Liu, W. Burgard, “Socially Compliant Navigation through Raw Depth Inputs with Generative Adversarial Imitation Learning,” IEEE International Conference on Robotics and Automation, arXiv: 1710.02543 [cs], Oct. 2017, arXiv: 1710.02543.

[2] Y.F. Chen, M. Liu, M. Everett, Jonathan P. How, “Decentralized Non-communicating Multiagent Collision Avoidance with Deep Reinforcement Learning,” IEEE International Conference on Robotics and Automation, pp. 285-292, 2017

[3]S. Kuutti, S. Jallah, R. Bowden, “Training Adversarial Agents to Exploit Weaknesses in Deep Control Policies,” IEEE International Conference on Robotics and Automation, pp. 108-114, 2020

[4] C. Chen, Y. Liu, S. Kreiss, A. Alahi, “Crowd-Robot Interaction: Crowd-Aware Robot Navigation With Attention-Based Deep Reinforcement Learning,” IEEE International Conference on Robotics and Automation, pp. 6015-6022, 2019

[5] T. Haarnoja, A. Zhou, P. Abbeel, S. Levine, “Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor,” Proceedings of the 35th International Conference on Machine Learning, PMLR 80:1861-1870, 2018.

[6] Chen. et al, CrowdNav, 2019, GitHub repository, https://github.com/vita-epfl/CrowdNav