기술

웨이스가 학습시킨 휴머노이드는 넘어진 횟수가 70번에서 61번으로 줄면 정말 나아진 것인가

같은 로봇을 두 번 시험하자 91번과 75번이 나왔습니다

방동걸 (주)웨이스 대표이사
2026년 11월 1일

가상 환경에서 두 발 휴머노이드를 걷게 학습시키면서, 걷는 도중 몸을 밀었을 때 몇 번 넘어지는지 500번씩 시험해 확인했습니다. 학습 전 70번이던 넘어짐은 두 번째 학습 뒤 61번으로 줄었습니다. 그런데 학습을 더 이어 간 뒤 같은 로봇을 두 번 시험하자 한 번은 91번, 한 번은 75번이 나왔습니다. 로봇은 그대로인데 결과가 16번 차이 났습니다.

제조 AI 운영체제 VEXPLOR를 만드는 웨이스는 로봇 AI 제어를 연구하면서 이 실험을 하고 있습니다. 여기서 웨이스가 먼저 배운 것은 걷는 법이 아니라 결과를 읽는 법이었습니다. 학습이 효과가 있었는지는 숫자 하나로 정해지지 않습니다. 같은 로봇을 여러 번 시험했을 때 결과가 얼마나 흔들리는지 먼저 알아야 판단할 수 있습니다. 이 글은 실험 기록에서 그 흔들림을 어디서 만났고 어떻게 다루었는지 정리한 것입니다.

01

웨이스는 가상 환경에서 두 발 로봇이 밀려도 넘어지지 않게 학습시키고 있습니다

실험은 GPU 컴퓨터 한 대(엔비디아 DGX Spark)에서 물리 시뮬레이션으로 진행합니다. 가상 세계 4,096개를 동시에 띄우고, 각 세계에서 같은 로봇이 정해진 거리만큼 걸어가 10초 동안 서 있기까지 제한 시간 안에 마치면 성공으로 봅니다. 사람이 규칙으로 짠 걷기 제어기(이하 교사)가 먼저 걷고, 신경망이 교사의 동작을 따라 배웁니다(모방학습). 그 뒤 강화학습으로 다듬는 방법도 시험했습니다.

항목	값	비고
동시 가상 세계	4,096개	GPU 1대
평소 걷기 시험	1,000번	회차마다
밀기 시험	500번	회차마다
신경망 성공률	0% → 98.6~99.2%	평소 걷기

평소 걷기는 1,000번 중 990번 안팎을 성공하는 수준까지 왔습니다. 평소 걷기에서 먼저 알게 된 것은 학습 방법을 바꿔도 실패율이 비슷한 자리에서 멈춘다는 점이었습니다.

방법	실패율	남은 실패
교사 혼자	0.93%	4걸음 뒤 시간 초과
모방학습 신경망(최고)	1.4%	4걸음 뒤 시간 초과
강화학습 8번 평균	1.4%	4걸음 뒤 시간 초과

남은 실패는 거의 전부 4걸음을 걸은 뒤 거리를 조금 못 채우고 시간을 넘기는 경우였고, 교사도 같은 비율로 같은 실패를 냈습니다. 학습 방법끼리 실패율을 견주기 전에, 그 차이가 시험의 흔들림보다 큰지부터 봐야 했습니다.

남은 과제는 걷는 도중 밀려도 넘어지지 않는 것입니다. 이 과제를 학습시키면서 회차마다 밀기 시험 500번의 넘어짐을 기록했습니다.

02

같은 로봇을 두 번 시험하자 넘어진 횟수가 91번과 75번으로 나왔습니다

밀렸을 때 버티도록 강화학습을 6회차까지 이어 간 결과는 아래와 같습니다. 회차마다 학습한 뒤 같은 조건으로 500번을 시험했습니다.

회차	밀기 500번 중 넘어짐	비율
학습 전	70번	14.0%
2회차	61번	12.2%
4회차	81번	16.2%
6회차	91번	18.2%

2회차만 보면 학습이 효과가 있는 것처럼 보이고, 6회차만 보면 학습이 로봇을 망친 것처럼 보입니다. 판단하기 전에 6회차 로봇을 그대로 두고 한 번 더 시험했더니 넘어짐은 75번이었습니다. 같은 로봇이 91번과 75번 사이를 오간 셈이니, 500번 시험에서 위아래로 15번 안팎(합쳐 30번 폭)의 오르내림은 학습 효과가 아니라 시험 자체의 흔들림으로 읽어야 했습니다. 그 폭으로 보면 61번에서 91번까지 6회차의 오르내림은 모두 제자리를 맴돈 것입니다.



같은 로봇을 두 번 시험해도 넘어진 횟수는 91번과 75번으로 달랐습니다

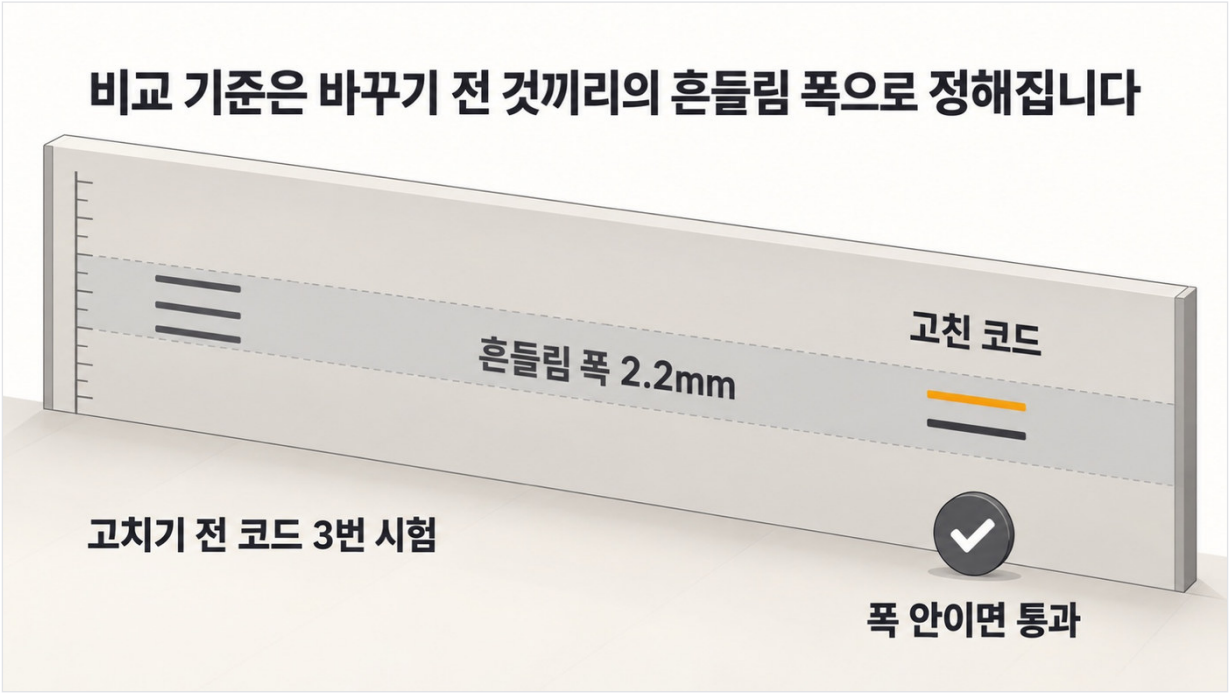
학습 기록을 봐도 마찬가지였습니다. 평가를 돕는 신경망의 오차는 2.23에서 1.74로 줄었지만, 로봇의 행동을 정하는 신경망은 거의 움직이지 않았습니다. 행동에 섞는 무작위 폭은 작게 고정돼 새로운 동작을 시도해 볼 여지가 적었습니다. 그 폭과 학습 속도를 바꿔 다시 해 봐도 학습 뒤 로봇이 출발점보다 나아진 적이 없어, 강화학습으로 다듬는 방법은 네 번 시도한 끝에 접고 모방학습으로 돌아갔습니다. 숫자의 오르내림을 학습 효과로 읽었다면 이 결정을 훨씬 늦게 내렸을 것입니다.

03

흔들림은 로봇만이 아니라 시험하는 스크립트와 고친 코드에서도 생깁니다

흔들림은 로봇에서만 생기지 않았습니다. 교사 제어기를 1,000번 시험한 결과가 한 번은 시간 초과 24번, 다른 한 번은 12번으로 나왔습니다. 처음에는 교사도 흔들린다고 봤지만, 다시 시험하자 12번이 그대로 나왔고 24번은 시험 스크립트가 달라서 생긴 차이로 확인됐습니다. 같은 로봇이라도 시험하는 방법이 바뀌면 숫자가 바뀝니다.

교사 코드를 고칠 때도 같은 문제가 있었습니다. 새 기능을 넣고 기능을 끈 상태로 64번 시험해 고치기 전과 같은 결과가 나오는지 확인했는데, 골반이 나아간 거리가 2mm 달라 미리 정한 기준(1mm)에 걸려 멈췄습니다. 고친 코드가 문제인지 보려고 고치기 전 코드를 두 번 더 시험해 모두 세 번의 결과를 보니, 같은 코드끼리도 거리가 0.1777m에서 0.1799m까지 2.2mm 흔들렸습니다. 고친 코드의 결과 0.1779m와 0.1785m는 그 폭 안에 있었고, 그제야 통과로 판정했습니다.



고치기 전 코드끼리의 흔들림 폭을 먼저 확인하고 그 폭으로 고친 코드를 판정하는 순서

이 일에서 얻은 기준은 하나입니다. **비교 기준은 머릿속으로 정하지 않고, 바꾸기 전 것끼리 벌어지는 폭을 먼저 확인해서 정합니다.** 1mm라는 기준은 그럴듯해 보였지만 실제 흔들림보다 좁았습니다.

04

어려운 장면을 3배로 반영한 실험에서 평소 걷기 성공이 959번에서 867번으로 떨어졌습니다

흔들림 폭을 알고 나면 실패도 빨리 알아볼 수 있습니다. 교사 제어기의 설정값 9가지와 코드 변경 2가지를 차례로 시험했지만, 기준 교사보다 나은 것이 하나도 없어 모두 탈락시켰습니다. 다음으로 신경망 쪽을 바꿔, 로봇이 거리를 다 못 가고 시간을 넘긴 장면과 도착한 뒤 디딤이 흔들린 장면의 학습 데이터를 3배로 반영하는 방법을 시험했습니다.

시점	평소 1,000번 중 성공	어려운 장면 비중
출발 로봇	959번	2.2%
1회차 뒤	867번	12.2%
2회차 뒤	938번	4.1%

1회차 뒤 성공이 92번 떨어졌습니다. 흔들림 폭을 훨씬 넘는 차이라 우연으로 볼 수 없었습니다. 학습 데이터에서 어려운 장면의 비중도 2.2%에서 12.2%로 커졌습니다. 어려운 장면을 더 배운 로봇이 그 장면을

더 많이 만들고, 그것이 다시 3배로 반영되는 되먹임이 원인 후보 가운데 하나입니다. 이 실험은 시작 전에 멈춤 기준(평소 성공이 3회차 연속 963번 미만이거나 밀기 넘어짐이 2회차 연속 20번 이상)을 적어 두었고, 1회차 결과를 본 뒤 「평소 성공 940번 미만」을 하나 더했습니다. 2회차가 938번으로 이 기준에 걸렸고, 마지막 로봇을 한 번 더 시험해도 945번에 그쳐 실험을 접었습니다. 지금은 반영 배수를 2배로 낮춰 다시 시험하고 있습니다. 출발 로봇을 세 번 시험한 폭은 954번에서 962번이었는데 1회차 뒤 906번이 나와, 이 폭 아래로 떨어진 원인을 더 좁히고 있습니다.

05

300번을 두 번 모두 성공해야 하는 판정은 99% 로봇에게 거의 열리지 않습니다

결과의 흔들림은 합격 기준을 정할 때도 따져야 합니다. 이 실험은 10월 중순까지 학습에 쓰지 않은 조건 300가지에서 두 번 연속 모두 성공해야 합격으로 판정했습니다. 밀기 없이 평지를 걷던 10월 18일 시험 결과는 300번 중 294번, 292번이었습니다. 거의 다 성공했지만 불합격입니다.

이 판정은 확률의 문제입니다. 시험 한 번의 성공 확률이 매번 같고 서로 영향을 주지 않는다고 보면, 600번을 모두 성공할 확률은 아래와 같습니다.

로봇의 실제 성공률	600번 모두 성공할 확률	뜻
99.9%	약 55%	두 번에 한 번꼴
99.5%	약 5%	스무 번에 한 번꼴
99.0%	약 0.2%	거의 불가능

1,000번 중 990번을 성공하는 로봇은 이 판정을 사실상 통과할 수 없습니다. 문제는 기준이 엄격하다는 데 있지 않습니다. 기준이 요구하는 실패율(약 0.1%)과 지금 로봇의 실패율(약 1%) 사이에 열 배 차이가 있다는 점을 숫자로 알고 시작해야 합니다. 이 차이를 모르고 학습만 반복하면 통과하지 못하는 이유를 계속 학습 방법에서 찾게 됩니다. 이 계산을 본 뒤 판정은 「학습에 쓰지 않은 6,000번에서 실패 18번 이하, 넘어짐 6번 이하」로 바뀌었습니다. 지금 수준에서도 나아졌는지 아닌지를 가를 수 있는 기준입니다.

06

로봇 학습 결과를 믿기 전에 다섯 가지를 먼저 확인합니다

이번 실험에서 정리한 확인 순서는 아래와 같습니다. 로봇이 아니라 공장 설비의 AI 제어나 검사 모델에도 같은 방식으로 쓸 수 있습니다.

1. **같은 대상을 세 번 이상 시험했는가** — 학습 전 로봇이나 고치기 전 코드를 여러 번 시험해 결과가 벌어지는 폭을 먼저 적습니다.
2. **차이가 그 폭을 넘는가** — 폭 안의 오르내림은 효과로 읽지 않습니다. 이번 실험에서는 500번 시험에서 위아래 15번 안팎이었습니다.
3. **시험 방법이 같은가** — 스크립트·조건·시험 횟수가 바뀌었으면 숫자를 바로 비교하지 않습니다.
4. **멈출 기준을 숫자로 적어 두었는가** — 「3회차 연속 963번 미만이면 멈춘다」처럼 시작 전에 적고, 도중에 더한 기준은 더한 시점을 함께 적습니다. 그래야 나빠지는 실험을 오래 끌지 않습니다.
5. **합격 기준이 지금 수준에서 열리는가** — 요구하는 실패율과 지금 실패율의 차이를 확률로 따져 보고 시작합니다.



이 글을 한 장으로 요약하면 이렇습니다

공장에서 AI가 설비를 움직이려면 그 판단을 현장에 보내기 전에 가상 환경에서 먼저 검증해야 합니다. 그런데 이번 실험에서 보듯 가상 검증의 결과도 흔들립니다. 검증을 믿을 수 있게 하는 것은 시험 횟수가 아니라, 같은 것을 다시 시험했을 때 얼마나 흔들리는지 먼저 알고 그 폭으로 결과를 읽는 습관입니다.

WACE(웨이스)는 공장이 스스로 판단하고 멈추지 않게 하는 제조 AI 운영체제 VEXPLOR를 만드는 회사다.

07

함께 읽어보세요

- AI가 PLC에 명령을 보내기 전에 같은 스캔으로 움직이는 가상 공장에서 먼저 실행해야 하는 이유
- 지멘스 연구진의 AI 에이전트는 가상 공장의 생산 과제를 93% 풀었습니다
- 지멘스는 공장 자동화율을 어떻게 80%까지 올리려 하는가

원문 blog.vexplor.com/post/humanoid-walking-test-noise
발행 2026년 11월 1일 · 방동걸 (주)웨이스 대표이사

회사 주식회사 웨이스 (WACE Inc.) · 대표이사 방동걸 · 사업자등록번호 439-87-02852
주소 (18462) 경기 화성시 동탄구 동탄영천로 150, 에이동(A동) 1702·1706호 (영천동, 실리콘앨리동탄)
연락처 전화 070-4730-0629 · contact@wace.me
홈페이지 wace.me · linkedin.com/company/wace-inc

원문 홈페이지 링크드인