maum.ai
Physical AI

CostNav: A Navigation Benchmark for Cost-Aware Evaluation of Embodied Agents

Company & Speaker Overview

Founded in 2014 and listed on KOSDAQ in 2021, MAUM.AI is widely regarded as a frontrunner in commercializing robotics foundational models across brownfield industries.

Korea's Leading Physical AI Pioneer maum.ai

2025

- Founding member & inaugural chair of the Physical AI Association
- Officially registered as a member of Qualcomm's IoT Accelerator Program
- CES 2025: Co-exhibited in collaboration with Qualcomm
- WoRV (Physical AI Technology) accepted to ICRA; invited Oral Presentation
- 2025 AI Excellence Summit — AI Technology Innovation Award

2024

- CES 2024: Recognized as a leading company in the CTA Trend Report
- Delivered a large-scale AI docent for the Gangneung Winter Olympics commemorative metaverse exhibition
- 2024 CICON Awards: Winner, Integrated AI Business category
- AI Defense Innovation Forum: Grand Prize recipient
- Korea's Robot Company of the Year 2024 (selected)
- WoRV (Physical AI Technology): Outstanding Paper Award, NeurIPS 2024

Haebin Seong — Senior Research Scientist, Self-Driving & Robotics Division, MAUM.AI

- Leading Project CostNav : A Navigation Benchmark for Cost-Aware Evaluation of Embodied Agents. CostNav shifts the evaluation paradigm from task success rates to business-relevant profit per run, modeling the complete economic lifecycle (CapEx and OpEx, including collision damage) and factoring in Service-Level Agreement (SLA) compliance for revenue generation. Slated for presentation at CES 2026.

- Key Manager of Data Collection Pipeline for robotics foundational models: Architect of End-to-End data pipeline including data operator management, data collection environment setup, data collection platform development, data processing, and data visualization, showing effective transfer of desktop agents to robotic embodied agents in project D2E, which ranked #2 in Huggingface Papers of the Day.

- Leading Project Physical AI data factory : Constructing Bi-Manual Humanoid Teleoperation Data Factory for automating physical human labor in manufacturing and household tasks. Joint collaboration project with Crowdworks AI and Scale AI.

Next GPT Moment is coming — Physical AI

The Physical AI market is vast—capturing even a small share can generate significant profits.
With earlier-than-expected revenue and greater industrial reach than ChatGPT, moving now secures share and asymmetric upside.

Morgan Stanley

“Robotics and Physical AI will likely have revenue impact earlier than expected”

Morgan Stanley via Investing.com
(post-GTC market analysis), Mar 2025

Investing.com

Why robotics and physical AI will likely have revenue impact earlier than expected

Investing.com | Author Navamya Acharya | Stock Markets

Despite skepticism stemming from past AI investments—such as autonomous driving in 2018, which has yet to yield widespread commercial success—historical patterns suggest semiconductor companies can generate revenue early in the investment phase.

“Most companies will make money when autonomous cars, or robots, are deployed, but we would argue that processor revenues can come earlier, as their hardware dominates the profits during that investment phase,” Morgan Stanley added

Physical AI Market Measured in Tens of Trillions of Dollars

First $10T market in human history

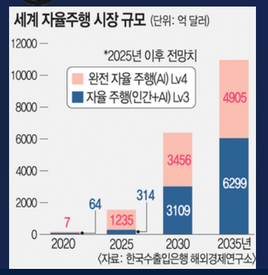
**Morgan Stanley prediction*

2035 Industrial Robot Market: $4T

**Roots Analysis prediction*

Autonomous Mobility Market

2035 Autonomous Driving Market: $1.1T
➡ 723% Growth



Even a small slice of the market can generate high profits

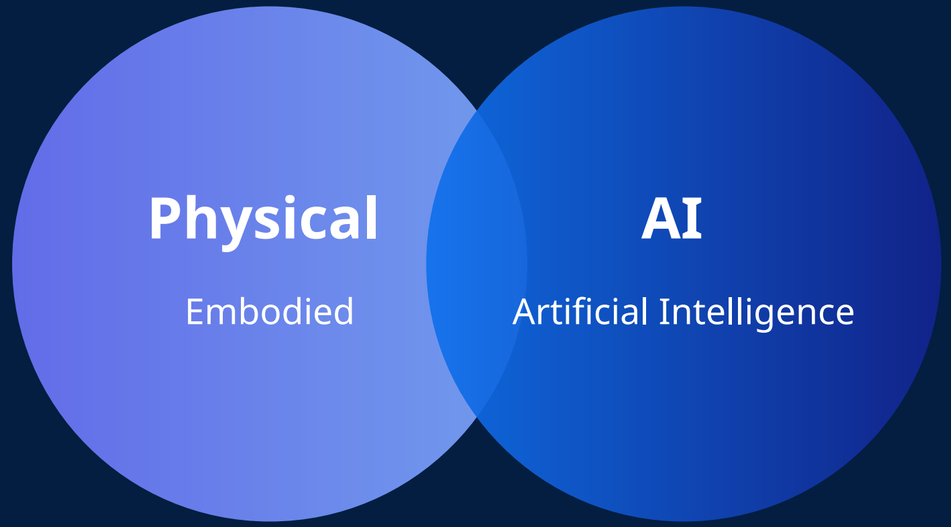
**Korea Eximbank prediction*

What is Physical AI?

New paradigm where AI systems have physical embodiment, interact with real-world environments, and perform physical human labor

---

Space-aware, real-time adaptive, environmentally coupled—AI built for the physical world



Data-driven:
How Physical AI Differs from Legacy Automation

Legacy Rule-Based Robots/Autonomy

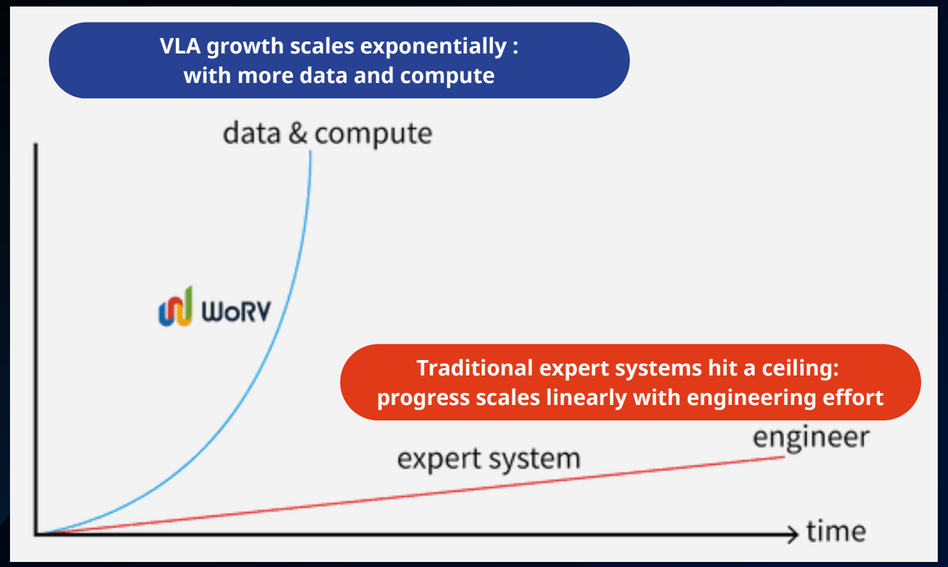
Hard-coded if–else rules

“If obstacle < 1m ahead and > 30cm high, turn right 90°, detour 2m, go straight.”

Physical AI Robots/Autonomy

Data-learned policies & VLA-driven planning

“Assess slope, crop spacing, and surface; avoid narrow aisles; slow down in mud.”



Rule-based

+ Deterministic and predictable
+ Relatively low compute and latency
+ Strong performance in controlled environment
- Brittle to edge cases
- Scales poorly with engineering effort
- Slow/expensive to retarget for new tasks

Data-driven

+ Adapts to variability/unstructured environments
+ Performance scales with data and compute
+ Rich perception/planning; handles noise better
- Higher upfront complexity and cost
- Less transparent/stochastic policies

The Broad Spectrum of Physical AI Across Industries

Physical AI attaches to existing industrial hardware to deliver human-level autonomy and immediate value across brownfield industries.

Physical AI is general-purpose—any machine that senses, decides, and acts can create new business value.

Agriculture

Autonomous Tractor
Speed Sprayer Robot

Logistics

Last Mile Delivery Robot
Long-haul Trailer Truck
Yard Truck / Tractor

Defense

Reconnaissance Robot
Tactical Support UGV

Construction

Autonomous Heavy Equipment
AGV/AMR Material Transport

Manufacturing

AMR
Welding Robot

Motivation

But is Physical AI ready for business?

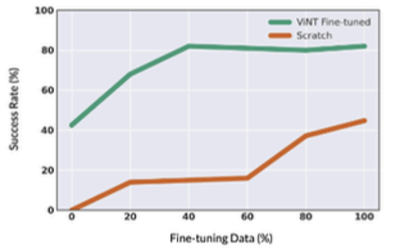
How can we deploy robots and make profit?

The Question no benchmark is answering

> **Traditional benchmarks only care about Navigation Success Rate**, which lacks context on how much **business value** it contains: the **cost** needed to implement the method, and the **revenue** generated by the task.

	Indoor: Position		Outdoor: GPS			Outdoor: Satellite		
Method	Success	Distance	Success	SPL	Distance	Success	SPL	Distance
ViKiNG [29]	0.60	56m	0.64	0.42	720m	0.77	0.68	780m
ViNT	**0.90**	**91m**	**0.95**	**0.84**	**1270m**	**1.00**	**0.94**	**1040m**

	Images		Positions	Routing
Method	Success	In Lane	Success	Success
Scratch	0.45	0.74	0.79	0.43
ImageNet	0.22	0.71	0.59	0.45
SimCLR [7]	0.21	0.63	0.70	0.64
VC-1 [44]	0.19	0.65	0.49	0.38
GNM [19]	0.49	0.66	0.45	0.49
ViNT	**0.82**	**0.82**	**0.89**	**0.72**



		Exploration		Navigation
Method	Params	Success	Coll.	Success
Masked ViNT^m	15M	50%	1.0	30%
VIB [17]	6M	30%	4.0	15%
Autoregressive^m	19M	90%	2.0	60%
Random Subgoals [3]	30M	70%	2.7	**90%**
Subgoal Diffusion [3]	335M	77%	1.7	**90%**
NoMaD	19M	**98%**	**0.2**	**90%**

Method	Params	Undirected	Goal-Conditioned
Diffusion Policy [31]	15M	98%	✗
ViNT Policy [3]	16M	✗	92%
NoMaD	19M	98%	92%

Will it make money?

CostNav: A Navigation Benchmark for Cost-Aware Evaluation of Embodied Agents

> **CostNav shifts the evaluation paradigm** from task success rates to **business-relevant profit** per run, modeling the **complete economic lifecycle** (CapEx and OpEx, including collision damage) and factoring in **Service-Level Agreement (SLA) compliance** for revenue generation.

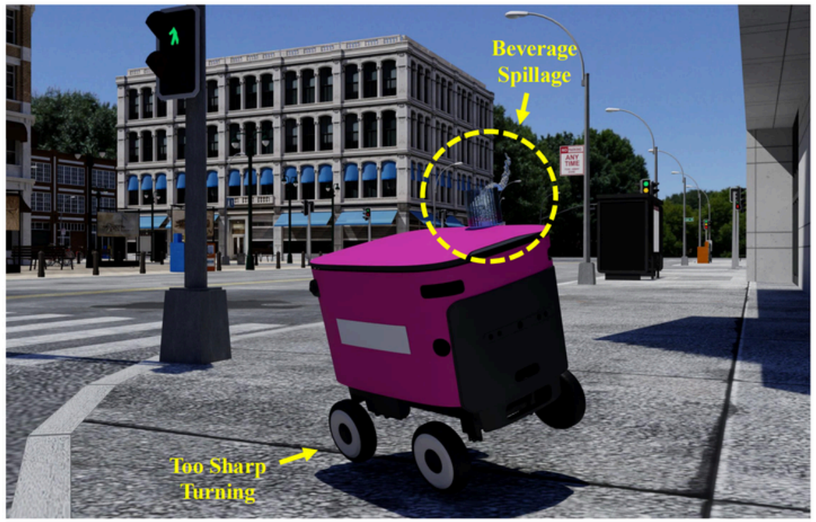


Figure 1. **A motivational example highlighting the core idea behind the CostNav benchmark.** Traditional metrics like success rate or collision rate overlook navigation behaviors that can lead to costly outcomes. For instance, overly sharp turning can spill beverages and cause unnecessary expenses. This gap motivates CostNav, which evaluates navigation through an economic lens.

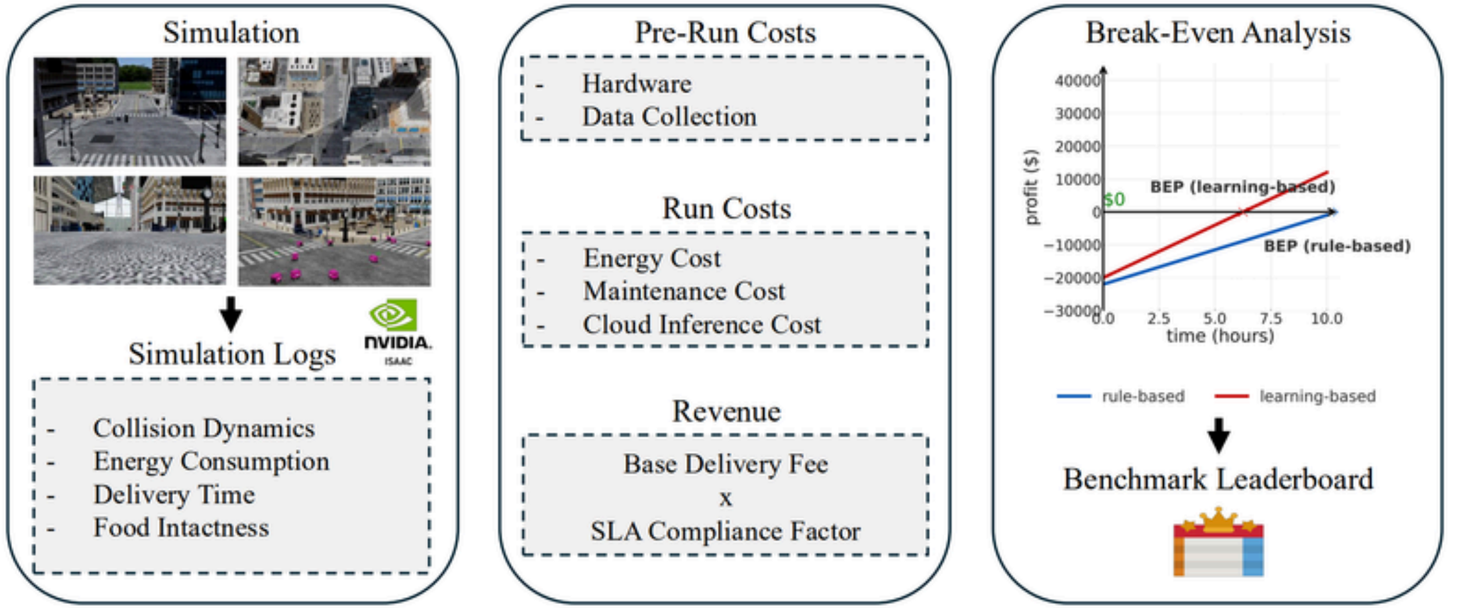


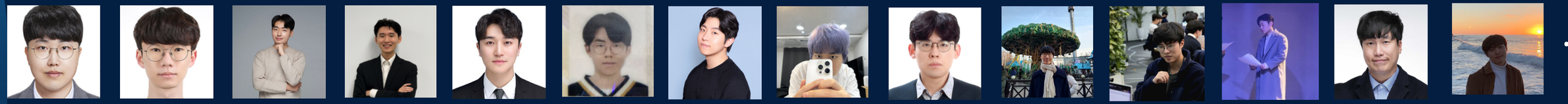
Figure 2. **End-to-end process of the CostNav benchmark, from simulation environments to break-even point analysis.** Simulation logs capture key operational signals—such as collision dynamics, energy usage, delivery time, and food intactness—that reflect how a robot behaves in realistic delivery scenarios. These signals are then combined with real-world cost and revenue models to compute profit curves and determine each method's break-even point. By translating navigation behaviors into economic outcomes, CostNav enables a leaderboard that ranks embodied agents based on financial performance rather than traditional task-centric metrics.

Teaser video

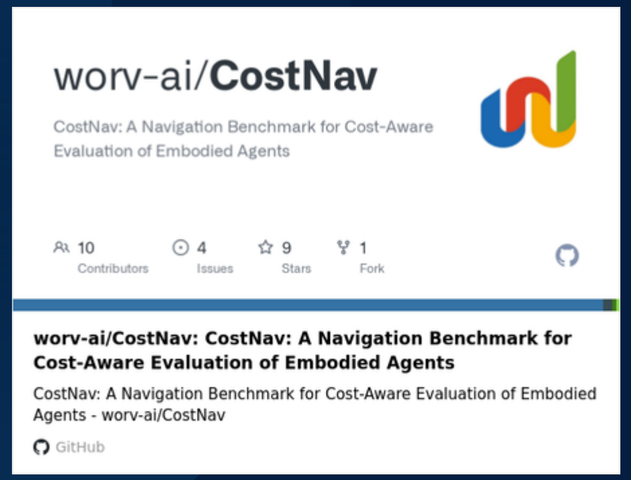
For the robot to be ready for food delivery business, it would require the robot to actually **deliver food without spoiling it.**

Open-source, Large-Scale, Realistic Benchmark

CostNav aims to be an **open-source large-scale public** benchmark with **numerous collaborators** inside & outside the company.
Our roadmap is to scale the benchmark with **continual updates** to be the de facto standard business benchmark that can actually give meaningful insights about **realworld physical AI value**.



Open-source github repo



https://github.com/worv-ai/CostNav

Blog post

[EN] CostNav: A Navigation Benchmark for Cost-Aware Evaluation of Embodied Agents

Hi, I'm Haebin Seong, Senior Researcher on the WoRV (World Model for Robotics and Vehicle Control) team at Maum.AI. I serve as the Project Lead for CostNav, one of the company's most strategically important...

WoRV Tech Blog / Nov 26

https://worv.ghost.io/costnav-2/

Large-Scale, Realistic

Maps: Urban Sidewalk, School, Hospital, Golf yard, Port, ...
Robots: Nova Carter, Segway, COCO, Spot, Go2, G1, ...
Scenarios: Weather, Lighting, Dynamic Traffic, ...
Methods: Nav2, Imitation Learning, RL, ViNT, SketchDrive, ...
Task: CostMani (Manipulation), CostMobileMani (Mobile Manipulation)

To be presented at CES 2026

CES™
Powered by The Consumer Technology Association®

We are looking for collaborators or teammates!

We are looking for people who are interested in collaborating in **CostNav**, or people who wants to **join our team**.
Feel free to reach out to us!

CostNav Keywords:

ROS, Isaac-Sim, Navigation,
Imitation Learning, RL,
financial consulting, accounting

Open Positions

LET US COME TOGETHER TO BUILD A UNIFIED ECOSYSTEM FOR PHYSICAL AI

— ONE THAT DEFINES THE GLOBAL STANDARD FOR THE NEXT ERA OF INTELLIGENCE