Embodied AI: Bridging Robotics and AI Toward Real-World Applications

By Hao Su, Yunduan Cui, Cosimo Della Santina, Kuan Fang, Jens Kober, Yanan Li,
Yunzhu Li, Takamitsu Matsubara, Maria Pozzi, and Bolei Zhou

Over the past decade, robotics has experienced substantial progress in perception, control, and learning. Advances in sensing, actuation, and computation have enabled robots to generate agile motions, process high-dimensional sensory inputs, and operate in increasingly complex environments. At the same time, it has become rapidly clear that sustained progress in robotics cannot be attributed to algorithms alone. Enabling robots to autonomously and reliably interact with humans and the physical world requires careful integration across sensing, perception, planning, control, and embodiment. Embodied artificial intelligence (AI) directly addresses this challenge by emphasizing closed-loop interaction, where intelligence emerges from the continuous coupling between physical bodies, learned representations, and real-world feedback.

In this special issue of *IEEE Robotics and Automation Magazine*, we adopt embodied AI not as a specific algorithmic approach but as a unifying perspective on how intelligence should be embedded into robotic systems. From this viewpoint, intelligence is shaped by physical embodiment, task constraints, sensing modalities, and interaction dynamics rather than existing as an abstract computational entity. Embodied AI therefore emphasizes system-level design, where learning, modeling, and control are codeveloped with hardware to achieve robust behavior in the physical world.

The articles collected in this special issue reflect the breadth of this perspective, spanning legged locomotion, dexterous manipulation, soft robotics, medical intervention, assistive systems, autonomous navigation, and long-horizon skill learning. Two survey contributions provide complementary overviews of perceptive legged locomotion in real-world environments and of dexterous and embodied manipulation, respectively [A1], [A2]. Together, they highlight recurring challenges related to robustness, physical interaction, sensing, and embodiment that motivate many of the technical contributions in the issue.

A central theme across the issue is the interdependence between hardware and AI. Physical embodiment shapes what information can be sensed, how forces are generated, and which behaviors are feasible. Learning, in turn, determines how effectively a system can exploit these physical capabilities. Rather than viewing intelligence as independent of embodiment, the contributions in this issue illustrate how mechanical design, actuation, sensing, and control objectives fundamentally influence learning outcomes. Examples include the use of compliance and morphology to simplify manipulation [A3], physically informed modeling for soft actuators [A9], and modular robotic platforms designed to support embodied learning across different morphologies [A5].

Embodied AI is inherently interdisciplinary and system driven. While learning-based methods have become increasingly powerful, integrating perception, learning, and control into reliable robotic systems remains a major challenge. Tasks involving dexterous manipulation, contact-rich interaction, and human–robot collaboration continue to demand careful consideration of uncertainty, safety, and physical constraints. Historically, learning-based approaches, e.g., deep reinforcement learning, were adopted cautiously in robotics due to concerns about sample efficiency and safety. Over time, advances in simulation, objective design, and structured policies have demonstrated that learning can be a practical component of embodied systems when grounded in physical models and constraints [1], [2], [3].

Sim-to-real learning is also becoming increasingly practical in domains involving physical interaction with humans. In medical robotics, high-fidelity digital twins and simulators enable the development and validation of autonomous behaviors that interact with tissue and internal anatomical structures prior to deployment, as demonstrated by the cardiac intervention system presented in this special issue [A4]. In wearable robotics, experiment-free learning in simulation has enabled effective exoskeleton assistance without requiring extensive human subject data collection, demonstrating both safety and practicality [4]. These advances highlight how simulation, when combined with embodiment-aware objectives and physical constraints, can serve as a powerful and pragmatic tool for embodied learning.

A recurring message across the special issue is that effective embodied intelligence does not require overly complex models or massive neural networks. Instead, well-chosen state representations,

Digital Object Identifier 10.1109/MRA.2026.3653534
Date of current version: 19 March 2026

physically grounded objectives, and modest-sized models can achieve robust real-world performance. This observation aligns with broader trends in physical intelligence and model-based reinforcement learning, where structure, embodiment, and feedback loops play a central role in shaping behavior and reducing unnecessary complexity [A10], [A11], [A12], [A13], [5], [8]. Data-driven learning is most effective when paired with appropriate physical priors, models, and constraints rather than used in isolation.

Large language models and foundation-model ideas are beginning to influence embodied systems by enabling new forms of abstraction, interaction, and task specification. In this special issue, such models appear primarily as high-level components that complement embodiment-aware perception, control, and planning. Examples include language-guided interactive navigation [A7] and language-assisted teleoperation [A6], where physical execution remains grounded in robot dynamics and sensing. These contributions underscore that data-driven learning alone is insufficient for reliable physical interaction; accurate models, physical priors, and constraint-aware control remain essential for safety, efficiency, and real-time responsiveness [6], [7].

Looking forward, embodied AI research faces several open challenges, including principled hardware–AI codesign, improved generalization across tasks and embodiments, and systematic evaluation of safety and robustness in real-world environments. Addressing these challenges will require continued collaboration across robotics, AI, and related disciplines.

Ultimately, the goal of embodied AI is not to develop algorithms in isolation but to create robotic systems that provide tangible benefits to people. Improving quality of life, enhancing safety, enabling medical interventions, and assisting humans in physically demanding or hazardous environments are the benchmarks by which success should be measured. The contributions in this special issue collectively reflect this goal, emphasizing embodiment, integration, and real-world relevance as defining characteristics of embodied AI research.

AUTHORS

Hao Su, Tandon School of Engineering, Courant Institute School of Mathematics, Computing, and Data Science, New York University, New York, NY 11201 USA. E-mail: hao.su@nyu.edu.

Yunduan Cui, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China. E-mail: yd.cui@siat.ac.cn.

Cosimo Della Santina, Delft University of Technology, Delft, The Netherlands. E-mail: C.DellaSantina@tudelft.nl.

Kuan Fang, Cornell University, Ithaca, NY 14853 USA. E-mail: kuanfang@cornell.edu.

Jens Kober, Delft University of Technology, Delft, The Netherlands. E-mail: j.kober@tudelft.nl.

Yanan Li, University of Sussex, Brighton, BN1 9RH, U.K. E-mail: yl557@sussex.ac.uk.

Yunzhu Li, Columbia University, New York, NY 10027 USA. E-mail: yunzhu.li@columbia.edu.

Takamitsu Matsubara, Nara Institute of Science and Technology, Ikoma City, Nara, 630-0192, Japan. E-mail: takam-m@is.naist.jp.

Maria Pozzi, University of Siena, 53100 Siena, Italy, and the Istituto Italiano di Tecnologia, Genoa, Italy. E-mail: maria.pozzi@unisi.it.

Bolei Zhou, University of California, Los Angeles, Los Angeles, CA, 90095, USA. E-mail: bolei@cs.ucla.edu.

APPENDIX: RELATED ARTICLES

[A1] I. T. Kurniawan, W. Zhu, D. Owaki, and M. Hayashibe, "Learning perceptive legged robot locomotion in the real world: A systematic review," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 8–23, Mar. 2026, doi: 10.1109/MRA.2025.3639794.

[A2] G. Li, R. Wang, P. Xu, Q. Ye, and J. Chen, "The developments and challenges toward dexterous and embodied robotic manipulation: A survey," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 24–38, Mar. 2026, doi: 10.1109/MRA.2025.3642671.

[A3] E. Turco et al., "Leveraging embodied mechanical intelligence for learning decluttering tasks: Gripper design boosts learning," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 39–51, Mar. 2026, doi: 10.1109/MRA.2025.3639825.

[A4] Y. Wang, M. Xu, W. Gaozhang, and H. A. Wurdemann, "From patient-specific digital twin to real-world phantom: Autonomous right heart catheterization," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 52–62, Mar. 2026, doi: 10.1109/MRA.2025.3642745.

[A5] O. Kwon et al., "PAPRLE: Plug-and-play robotic limb environment: A modular ecosystem for robotic limbs," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 63–73, Mar. 2026, doi: 10.1109/MRA.2025.3647246.

[A6] H. Fei et al., "Large-language-model-aided assistive robot for single-operator bimanual teleoperation: Introduction and validation of a flexible assistance system," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 74–84, Mar. 2026, doi: 10.1109/MRA.2025.3642669.

[A7] K. Zhou et al., "AINav: Large language model-based adaptive interactive navigation," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 85–100, Mar. 2026, doi: 10.1109/MRA.2025.3639793.

[A8] H. Xu, X. Chen, Y. Lang, and Q. Ren, "Vision-based policy learning for high-speed autonomous racing: A two-phase learning paradigm," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 101–114, Mar. 2026, doi: 10.1109/MRA.2025.3639812.

[A9] S.-Y. Lee, L. Z. Yañez, J. Rogatinsky, V. T. Vo, T. Shingade, and T. Ranzani, "Simplifying data-driven modeling of the volume-flow-pressure relationship in hydraulic soft robotic actuators: A practical and balanced solution," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 115–129, Mar. 2026, doi: 10.1109/MRA.2025.3642670.

[A10] Z. Ajanović, R. Prakash, L. de Souza Rosa, and J. Kober, "Sequentially teaching sequential tasks (ST)^2: Teaching robots long-horizon manipulation skills," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 130–141, Mar. 2026, doi: 10.1109/MRA.2026.3650853.

[A11] D. Dona', G. Franzese, C. Della Santina, P. Boscariol, and B. Lenzo, "Real-time generation of near minimum-energy trajectories via constraint-informed residual learning: A paradigm for learning from optimal solutions," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 142–150, Mar. 2026, doi: 10.1109/MRA.2025.3642672.

[A12] Y. Kadokawa, J. Frey, T. Miki, T. Matsubara, and M. Hutter, "DAPPER: Discriminability-aware policy-to-policy preference-based reinforcement learning for query-efficient robot skill acquisition," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 151–166, Mar. 2026, doi: 10.1109/MRA.2026.3650847.

[A13] Y. Yan and D. Lee, "Cross-embodiment imitation: Learning a unified latent space for multirobot control," *IEEE Robot. Autom. Mag.*, vol. 33, no. 1, pp. 167–179, Mar. 2026, doi: 10.1109/MRA.2026.3651673.

REFERENCES

[1] J. Hwangbo, J. Lee, S. Jeong, and M. Hutter, "Learning agile and dynamic motor skills for legged robots," *Sci. Robot.*, vol. 4, no. 26, 2019, doi: 10.1126/scirobotics.aau5872.

[2] D. Kalashnikov et al., "QT-Opt: Scalable deep reinforcement learning for vision-based robotic manipulation," 2018, *arXiv:1806.10293*.

[3] I. Akkaya et al., "Solving Rubik's cube with a robot hand," *Int. J. Rob. Res.*, 2019.

[4] S. Luo et al., "Experiment-free exoskeleton assistance via learning in simulation," *Nature*, vol. 630, no. 8016, pp. 353–359, 2024, doi: 10.1038/s41586-024-07382-4.

[5] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp 1334–1373, 2016.

[6] Z. Fu, T. Z. Zhao, and C. Finn, "Mobile ALOHA: Learning bimanual mobile manipulation with low-cost whole-body teleoperation," 2024, *arXiv:2401.02117*.

[7] A. Amin et al., "$\pi^{*}_{0.6}$: A VLA that learns from experience," 2025, arXiv:2511.14759.

[8] J. Kwon, K. Choi, and F. C. Park, "Kinodynamic model identification: A unified geometric approach," *IEEE Trans. Robot.*, vol. 37, no. 4, pp. 1100–1114, Aug. 2021, doi: 10.1109/TRO.2020.3047515.