SCIENCE CHINA
Information Sciences

• PERSPECTIVE •

November 2025, Vol. 68, Iss. 11, 216201:1–216201:3
https://doi.org/10.1007/s11432-025-4532-0

Humanoid robots: progress, challenges, and future research directions

Wenjuan LI[1,2], Jiyi WU[1,3*], Genyuan YANG[2], Ben WANG[2], Qifei ZHANG[4], Yuqi WU[3] & Jianrong TAN[1]

[1]*Intelligent Education Research Center, Zhejiang University, Hangzhou 310058, China*
[2]*School of Information Science and Technology, Hangzhou Normal University, Hangzhou 311121, China*
[3]*Advanced Institute of Information Technology, Peking University, Hangzhou 310015, China*
[4]*School of Software Engineering, Zhejiang University, Ningbo 315048, China*

Received 22 March 2025/Revised 9 June 2025/Accepted 14 July 2025/Published online 8 September 2025

Citation Li W J, Wu J Y, Yang G Y, et al. Humanoid robots: progress, challenges, and future research directions. Sci China Inf Sci, 2025, 68(11): 216201, https://doi.org/10.1007/s11432-025-4532-0

Humanoid robots are autonomous systems with human-like forms and intelligent decision-making capabilities, enabling them to interact naturally with humans in complex environments [1]. Compared with traditional industrial robots, humanoid robots have stronger environmental adaptability, operational flexibility, and human-robot interaction (HRI) abilities, making them capable of performing various tasks in human-centered environments [2]. The humanoid form factor enables direct integration into human-designed environments and facilitates intuitive human-robot interaction. Currently, the technological advances of representative products such as Tesla Optimus and Boston Dynamics Atlas mark the transition of humanoid robots from laboratory research to commercial applications [3]. Figure 1 illustrates the humanoid robot ecosystem.

Breakthroughs in deep learning, advanced materials, and computing capabilities have spurred the rapid development of humanoid robots [4]. In academic research, significant progress has been made in perception, control, and decision-making [1, 2]. Multi-sensor fusion and deep learning have enhanced environmental understanding and object recognition capabilities, motion control has evolved from traditional zero-moment point (ZMP) methods to deep reinforcement learning and bio-inspired approaches enabling more natural movement [5], and the integration of reinforcement learning, imitation learning, and multi-modal interaction has endowed robots with autonomous adaptation and human collaboration capabilities [6, 7]. The industry is actively positioning itself with continuously emerging products and rapidly growing markets. Representative products like Tesla Optimus, Boston Dynamics Atlas, and Unitree G1 showcase different technological approaches: Tesla focuses on general factory operations, Boston Dynamics excels in agile movement, and emerging Chinese robot companies have formed competitive advantages in cost control. The humanoid robot industry is transitioning from technological validation to scaled production.

Currently, humanoid robots are demonstrating potential applications in fields such as medical rehabilitation, the service industry, and industrial manufacturing. In the medical field, they assist rehabilitation training through emotional interaction [8]. In the service industry, the advantages of natural HRI have been verified. In industrial applications, companies like Tesla and Honda are testing the commercial feasibility of assembly and logistics tasks. The unique human-like form and environmental adaptability of humanoid robots offer new solutions for complex interaction scenarios that traditional robots find difficult to handle.

However, the research and application of humanoid robots are still in the early stages, facing numerous technical challenges and difficulties.

• Hardware system design. The hardware system is subject to dual constraints of performance and cost. Actuators need to meet requirements for high power density, high torque, and reversible drive, while the cost of a single joint is excessively high [2]. Electromagnetic interference causes data loss, and sensors face trade-offs between precision and shock resistance. The bottleneck in the energy system is even more prominent, with power consumption and endurance limitations as evidenced in current commercial prototypes [4].

• Control system design. The existing software architecture lacks a unified middleware standard. Real-time communication delays between perception, decision-making, and control modules affect dynamic response performance, and the synchronization and coordination mechanisms of distributed control nodes are not well-developed. Humanoid robots require a motion control system that simultaneously possesses environmental adaptability, flexibility, motion accuracy, and coordination [4]. In tasks such as stable walking on complex terrain, precise grasping, and dynamic balance, it is necessary to handle multi-constraint optimization problems, including joint limitations, self-collision avoidance, and task space control, with extremely high algorith-

* Corresponding author (email: cloudlab@139.com)

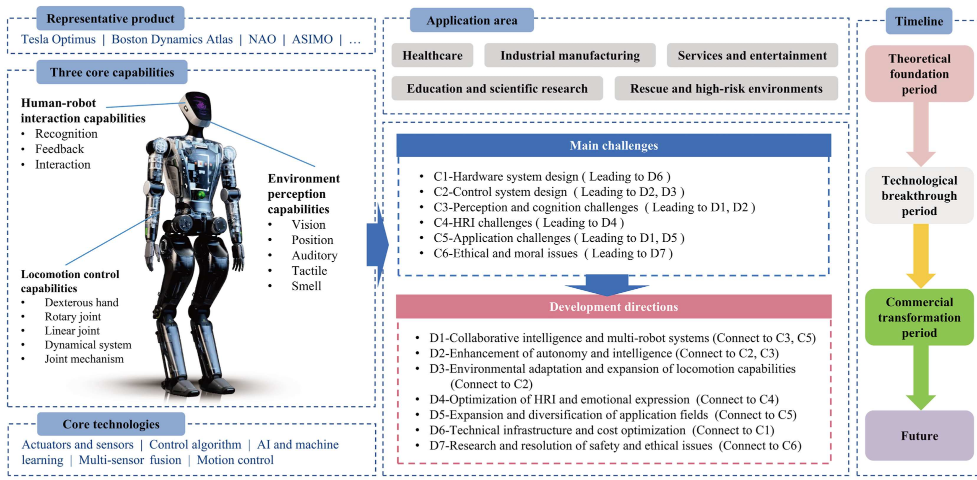

Figure 1 (Color online) Humanoid robot ecosystem.

mic complexity [5]. The motion control precision is insufficient to meet the demands of fine operations, and robustness under external disturbances is inadequate.

• Perception and cognition challenges. The capability for environmental perception is limited, with the core bottleneck being the physical contradiction between precision and coverage range. Multi-sensor information fusion faces issues such as significant differences in sampling frequencies, diverse data formats, and distinct feature spaces. Traditional fusion methods struggle to effectively handle the temporal and spatial alignment, as well as the feature correlation, of heterogeneous data [7]. Dynamic changes in complex environments severely affect the stability of perception systems. Although deep learning algorithms have significantly enhanced learning capabilities, their high computational complexity makes it difficult to meet real-time requirements on mobile platforms with limited computing resources. Moreover, robots lack semantic understanding of complex scenes and have deficiencies in common-sense reasoning and causal relationship analysis. Their decision accuracy drops significantly when facing unknown scenes, making it difficult to achieve true environmental generalization.

• HRI challenges. In complex scenarios, robots need to simultaneously process users' verbal instructions, facial expressions, body language, and emotional states. Current technologies have significant limitations in handling ambiguous expressions, metaphors, cultural differences, noisy environments, and multi-person settings, while accurately identifying user intentions remains challenging [7]. Moreover, although humanoid robots are becoming increasingly human-like in appearance, they have not yet achieved true "humanization". There are still deficiencies in the naturalness, consistency, and appropriateness of emotional expression. In specialized fields such as education and healthcare, robots need to develop personalized interaction strategies for different user groups, and establishing user models for long-term personalized adaptation is a significant challenge.

• Application challenges. Real-time visual processing still faces technological challenges in terms of accuracy, speed, and adaptability to complex environments. In high-risk environments such as disaster response, nuclear management, and mine rescue, robots need to operate stably under extreme temperature differences and radiation conditions, which pose higher demands on design and control. Their human-like form provides unique advantages in accessing confined spaces and collaborating with human teams. The design and manufacturing of humanoid robots are highly complex and costly, with actuators representing the dominant cost component in products like Tesla Optimus and Boston Dynamics Atlas [2], creating an unbalanced cost structure that limits large-scale production. The current production scale is only at the level of hundreds of units, with prohibitively high unit costs that limit large-scale production. Controlling the design and manufacturing costs of the products and improving the cost-performance ratio are key factors in promoting mass production.

• Ethical and moral issues. Humanoid robots face three major ethical challenges: algorithmic transparency, privacy protection, and attribution of responsibility. The "black box" nature of decision-making algorithms leads to insufficient explainability [9], and there is a contradiction between the protection of personal data privacy and service performance. The introduction of methods such as differential privacy significantly degrades the quality of interaction. Moreover, the lack of technically based mechanisms for tracing behavior and determining responsibility makes it difficult to accurately define liability and assess risks in the event of an accident.

In response to the above challenges, future research on humanoid robots can focus on the following areas.

• Collaborative intelligence and multi-robot systems. (1) Large-scale heterogeneous group collaboration: Develop multi-agent reinforcement learning frameworks for adaptive task allocation and millisecond-level coordination in large-scale heterogeneous groups. (2) Cognitive fusion and knowledge sharing: Construct a distributed cognitive architecture to enable knowledge sharing among multiple robots while protecting privacy, and utilize large language models (LLMs) to achieve group task understanding and decomposition based on natural language instructions. (3) Hybrid virtual-physical collaboration ecosystem: Build a hybrid virtual-physical collaboration framework based on digital twin technology, and progressively explore quantum-inspired optimization and neuromorphic computing architectures as these technologies mature, targeting future low-power real-time coordination for ultra-large-scale groups.

• Enhancement of autonomy and intelligence. (1) Innovation in autonomous learning and adaptation algorithms: Develop autonomous learning algorithms for complex environments, addressing exploration challenges in sparse reward settings. Integrate deep reinforcement learning, imitation learning, and transfer learning to construct a meta-learning framework for rapid task adaptation [6]. (2) Cognitive architecture and reasoning capability enhancement: Build a hierarchical decision-making architecture integrating working memory, attention selection, and executive control modules. Develop causal and common-sense reasoning algorithms to enhance autonomous decision-making in uncertain environments [4]. (3) Integration of embodied intelligence and language understanding: Integrate LLMs into robot control loops for end-to-end learning from natural language instructions to action sequences, supporting zero-shot task transfer and multimodal instruction understanding. (4) Neuromorphic computing and bionic intelligence: Progressively integrate neuromorphic chips for low-power perception processing, and explore spiking neural networks and bio-inspired learning mechanisms.

• Environmental adaptation and expansion of locomotion capabilities. (1) Perception and adaptation to complex terrain: Develop intelligent terrain recognition algorithms based on multi-sensor fusion to achieve real-time modeling of complex environments and predictive terrain perception, enabling autonomous navigation on complex terrains such as stairs, slopes, and obstacles. (2) Adaptive motion control technology: Develop variable stiffness actuators, adaptive gait control, and real-time footstep planning with force feedback for dynamic balance. (3) Bionic motion and morphological reconstruction: Develop bio-inspired control algorithms based on central pattern generators (CPG) and cerebellar adaptation models, and explore reconfigurable robots and hybrid soft-rigid technologies to support dynamic morphological adjustments according to environmental and task requirements.

• Optimization of HRI and emotional expression. (1) Intelligent dialogue and intent understanding: Develop a multi-turn dialogue system based on LLMs to break through intent recognition and contextual understanding in complex scenarios. Through multi-modal information fusion algorithms, achieve coordinated understanding of language, gestures, and facial expressions, supporting continuous and long-term natural dialogue. (2) Affective computing and coordinated expression: Construct an emotion recognition framework that integrates facial expressions, vocal prosody, and physiological signals. Develop a multi-modal emotional expression system to achieve natural and consistent emotional expression with cultural adaptability. (3) Personalized social interaction intelligence: Establish dynamic user models and long-term memory mechanisms to enable adaptive adjustment of personalized interaction strategies. Develop algorithms for understanding social norms and cultural sensitivity to enhance the appropriateness and acceptance of interactions with users from diverse backgrounds.

• Expansion and diversification of application fields. (1) Industrialization of intelligent services: Promote the large-scale commercial application of humanoid robots in service industries such as domestic services, education, healthcare, and entertainment, exploring the balance point of cost-effectiveness. (2) Specialized applications in extreme environments: Develop customized humanoid robots for high-risk environments such as disaster response, nuclear management, and mine rescue. Achieve technological breakthroughs in reliability under extreme conditions through modular redundant design and self-healing mechanisms. (3) Exploration of interdisciplinary innovative applications: Explore the potential applications of humanoid robots in emerging fields such as neuroscience, artistic creation, and sports training, driving breakthroughs in multimodal interaction, creative algorithms, and precision control technologies.

• Technical infrastructure and cost optimization. (1) Breakthroughs in emerging hardware technologies: Revolutionary technologies such as flexible electronic skin, soft robotics, and self-healing materials reshape the form and function of humanoid robots. Flexible electronic skin enables distributed sensing with shock resistance and tactile sensitivity, soft technologies enhance safe interaction, and biocompatible materials improve HRIs. The integration of these technologies significantly enhances the robots' environmental adaptability and the naturalness of interaction. (2) Cost optimization and industrialization promotion: Reduce the cost of actuators through integrated motor and reducer design, and develop high-efficiency drivers to significantly lower the overall cost of the robot. Optimize energy management strategies to extend battery life, laying the foundation for large-scale commercial deployment.

• Research and resolution of safety and ethical issues. Construct a multi-layered safety assurance system, including hardware safety constraints, software behavioral boundaries, and real-time monitoring mechanisms. Develop data processing frameworks that protect privacy and decision-making algorithms that are interpretable, and establish mechanisms for sharing responsibility between humans and robots and tracing accidents. Formulate ethical behavioral standards that are adaptable to different cultural contexts to ensure the safety of environments where humans and robots coexist [9].

We analyze humanoid robot technology from academic and industrial perspectives, examining current challenges and development trends. Humanoid robots are transitioning from laboratory to commercialization, with products like Tesla Optimus demonstrating feasibility but facing cost, stability, and adaptation challenges [1, 3, 4]. We identify technological bottlenecks and propose development directions, including collaborative intelligence integration. Widespread adoption requires addressing technological innovation and cost optimization.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61702151) and Joint Funds of the Zhejiang Provincial Natural Science Foundation of China (Grant No. LHZSZ24F020001).

References

1 Fukuda T, Dario P, Yang G Z. Humanoid robotics—history, current state of the art, and challenges. Sci Robot, 2017, 2: eaar4043
2 Saeedvand S, Jafari M, Aghdasi H S, et al. A comprehensive survey on humanoid robot development. Knowledge Eng Rev, 2019, 34: e20
3 Boston Dynamics. An electric new era for Atlas. https://bostondynamics.com/blog/electric-atlas/, 2024
4 Tong Y, Liu H, Zhang Z. Advancements in humanoid robots: a comprehensive review and future prospects. IEEE CAA J Autom Sin, 2024, 11: 301–328
5 Kuindersma S, Deits R, Fallon M, et al. Optimization-based locomotion planning, estimation, and control design for the Atlas humanoid robot. Auton Robot, 2016, 40: 429–455
6 Lobos-Tsunekawa K, Leiva F, Ruiz-del-Solar J. Visual navigation for biped humanoid robots using deep reinforcement learning. IEEE Robot Autom Lett, 2018, 3: 3247–3254
7 Su H, Qi W, Chen J, et al. Recent advancements in multimodal human C robot interaction. Front Neurorobot, 2023, 17: 1084000
8 Lau Y, Chee D G H, Chow X P, et al. Humanoid robot-assisted interventions among children with diabetes: A systematic scoping review. Int J Nurs Studies, 2020, 111: 103749
9 Torras C. Ethics of social robotics: Individual and societal concerns and opportunities. Annu Rev Control Robotics Autonomous Syst, 2024, 7: 1–18