**Вісник ЛДУБЖД**
**Bulletin of Lviv State University of Life Safety**
**https://journal.ldubgd.edu.ua/index.php/Visnuk**
**ISSN 2078-4643 (print), ISSN 2708-1389 (online)**

**UDC 004(06)**
**DOI: 10.32447/20784643.32.2025.06**


***D. V. Kotelovych, Yu. O. Borzov, N. Ye. Burak***
*Lviv State University of Life Safety, Lviv, Ukraine*
ORCID: https://orcid.org/0009−0007−7031−2500 – D. V. Kotelovych
https://orcid.org/0000-0002-0604-0498 – Yu. O. Borzov
https://orcid.org/0000-0002-3880-4077 – N. Ye. Burak
denys.kotelovych@gmail.com

# DYNAMIC STABILITY IN LEGGED ROBOTICS: ADVANCES AND CHALLENGES

Dynamic stability is a core requirement for legged robots operating under modeling error, terrain variability, and intermittent ground contact. This article reviews published approaches that contribute to stable locomotion and whole-body behavior, drawing on commercial, open-source, and academic platforms.

**Purpose**. To synthesize and align model-based and learning-based methods for stability with sim-to-real practices, state estimation, and hardware trends, and to highlight open challenges and actionable guidance for deployment.

**Methods**. We summarize model-based methods, including the Linear Inverted Pendulum Model, centroidal dynamics, and hierarchical inverse dynamics with quadratic programming, together with preview-horizon model predictive control. We then survey learning-based controllers such as PPO, SAC, DDPG, TD3, imitation learning (including adversarial and example-guided variants), hierarchical policies, meta-learning, and recent transformer-based policies for bipedal locomotion. We further review domain and dynamics randomization, curriculum design, policy distillation, progressive networks, and Rapid Motor Adaptation for transfer, and outline EKF/IEKF fusion of leg kinematics and inertial data, multiple-model/contact-aware filters, invariant neural-augmented Kalman filtering, and the use of low-rate vision/LiDAR updates in Pronto for estimation; finally, we note hardware trends including compliant and series-elastic actuation, energy-storing linkages, and thruster assistance.

**Results**. A reported MuJoCo study on a Unitree Go1 highlights a representative trade-off: preview-horizon MPC rejected larger pushes, while PPO achieved a lower cost of transport on flat ground. Transfer techniques (domain/dynamics randomization, curricula, distillation, progressive nets, RMA) improve robustness to terrain and payload changes and rapid gait transitions; estimation pipelines with leg–IMU EKF/IEKF, multiple-model/contact-aware filters, invariant neural-augmented Kalman filtering, and Pronto's low-rate vision/LiDAR updates bound long-term drift. Safety-aware control elements – constrained policy optimization, backup controllers, and Lyapunov/CBF-based layers – further reduce fall rates and hazardous actions.

**Conclusion**. The review identifies open challenges reported in the literature, including formal safety for learned policies, robustness under contact-mode uncertainty, and practical sim-to-real pipelines that maintain performance on hardware. We recommend integrating CBF/Lyapunov shields with learned controllers, standardizing energy-aware loco-manipulation benchmarks, expanding contact-aware/invariant-state estimators, and prioritizing compliant, energy-storing actuation to achieve safe, efficient, and generalizable real-world behavior.

**Keywords:** centroidal dynamics; hierarchical inverse dynamics; model predictive control; impedance control; reinforcement learning; imitation learning; sim-to-real transfer; domain and dynamics randomization; state estimation (EKF/IEKF).

***Д. В. Котелович, Ю. О. Борзов, Н. Є. Бурак***
*Львівський державний університет безпеки життєдіяльності, м. Львів*

# ДИНАМІЧНА СТІЙКІСТЬ КРОКУЮЧИХ РОБОТІВ: ВИКЛИКИ ТА ПЕРСПЕКТИВИ

Динамічна стійкість є базовою вимогою для крокуючих роботів, що працюють за умов похибок моделі, мінливого рельєфу та переривчастих контактів із поверхнею. У статті розглянуто опубліковані підходи, що сприяють стабільній ході та цілісній поведінці всього тіла, з урахуванням комерційних, відкритих і академічних платформ.

**Мета**. Узагальнити та зіставити модельно орієнтовані й навчальні методи забезпечення стійкості, практики перенесення із симуляції в “реальність”, підходи до оцінювання стану та релевантні апаратні рішення, а також виокремити відкриті виклики та пропозиції для подальших досліджень.

**Методи**. Узагальнено модельно орієнтовані методи, зокрема лінійну модель оберненого маятника, центроїдну динаміку, ієрархічну інверсну динаміку з квадратичним програмуванням та MPC з горизонтом передбачення. Підсумовано навчальні методи: PPO, SAC, DDPG, TD3, імітаційне навчання (включно із змагальним та прикладно керованим), ієрархічні політики, метанавчання та новітні політики на основі трансформерів для двоногого пересування. Розглянуто засоби перенесення (доменна й динамічна рандомізація, навчання за програмою, дистиляція політик, progressive nets, Rapid Motor Adaptation), оцінювання стану (EKF/IEKF-злиття кінематики ніг та інерціальних даних, мультимодельні/контакт-обізнані фільтри, інваріантний нейромережею доповнений фільтр Калмана, низькочастотні візуальні/LiDAR-оновлення у Pronto) та апаратні тренди (пружні й послідовно-еластичні приводи, енергонакопичувальні ланки, допоміжні рушії).

**Результати**. Наведене дослідження в MuJoCo на Unitree Go1 ілюструє типовий компроміс: MPC з передбачувальним горизонтом ефективніше компенсував зовнішні поштовхи великої амплітуди завдяки явній оптимізації із врахуванням обмежень, тоді як PPO забезпечив нижчу енергоємність руху на рівній поверхні (менший cost of transport). У частині перенесення показано, що доменна й динамічна рандомізація, змішане навчання за програмою та RMA підвищують надійність до змін рельєфу, навантаження і швидких переходів між типами ходи; дистиляція політик та progressive nets зменшують «забування» і спрощують багатозадачні політики. Для оцінювання стану підтверджено ефективність EKF/IEKF-інтеграції leg–IMU, мульти-модельних/контакт-обізнаних фільтрів та інваріантного нейро-доповненого Кальмана, а також користь низькочастотних візуальних/LiDAR-оновлень у Pronto для стримування довготривалої помилки. Коротко означено апаратні чинники стійкості: комплаєнтні та SEA-приводи, енергонакопичувальні ланки й допоміжні рушії, що розширюють робочу область стабільних режимів. Додатково узагальнено елементи безпеково-орієнтованого керування – оптимізацію політик з обмеженнями, резервні регулятори та шари на основі функцій Ляпунова/CBF – як механізми зниження частоти падінь і ризикованих дій.

**Висновки**. Стаття виокремлює відкриті виклики, зазначені в літературі: формальне забезпечення безпеки для навчальних політик, працездатність під час збоїв або втрат контакту та методики перенесення з симуляції на реальних роботів, що зберігають продуктивність на апаратній платформі. Практично значущими кроками є інтеграція CBF/Ляпунова як захисних шарів до RL-політик, стандартизація бенчмарків із енергетичними метриками та сценаріями локоманіпуляції, ширше застосування контакт-обізнаних/інваріантних оцінювачів стану та пріоритет апаратних конфігурацій із комплаєнтними приводами й енергоакумуляцією, що разом сприятиме безпечній, енергоефективній та узагальнюваній поведінці крокуючих роботів у реальному світі.

**Ключові слова**: центроїдна динаміка; ієрархічна інверсна динаміка; модельно-прогнозне керування (MPC); імпедансне керування; навчання з підкріпленням; імітаційне навчання; перенесення з симуляції; доменна й динамічна рандомізація; оцінювання стану (EKF/IEKF).

**Introduction.** We are living in an era of remarkable technological transformation, where breakthroughs in robotics, artificial intelligence, and computing are converging to redefine our vision of the future. The rapid progress in legged robots is paralleled by advancements in large language models, edge computing, advanced sensors, energy-efficient actuators, and bio-inspired designs. These innovations collectively propel robots from clumsy prototypes to increasingly versatile, human-like machines. Developments in machine learning, reinforcement learning, and sim-to-real transfer have dramatically improved robots’ ability to adapt and perform in dynamic environments, while hardware advances make them lighter, more agile, and power-efficient.

This convergence hints at the emergence of truly general-purpose robots – machines capable of understanding, interacting, and responding to the world with unprecedented sophistication. As we stand on the cusp of a future where human-robot collaboration is commonplace, it is a privilege to contribute to this fast-evolving field with the potential to reshape our lives.

At the heart of this transformation lies the challenge of dynamic stability, a fundamental capability. Dynamic stability is essential not only for stable locomotion but also for enabling manipulation and coordinated movements in dynamic and unpredictable environments, including human interaction [1], [2]. Ensuring dynamic stability requires managing complex dynamics and adapting continuously to changing conditions, challenge that forms the cornerstone for developing versatile, capable robots.

**Purpose.** This paper explores the state of the art in robotic dynamic stability by reviewing key developments from both industry and research perspectives. We begin by reviewing major commercial projects, then examine open-source and research initiatives. Following this industry review, we provide a literature survey that delves into model-based control, reinforcement learning approaches, sim-to-real transfer methods, state estimation, and hardware design innovations. By synthesizing insights from practical implementations and theoretical research, we identify key challenges and chart potential future directions for teaching robots to achieve dynamic stability effectively.

*Object of research* is the process of achieving dynamic stability in legged robots under diverse environmental conditions.

*Subject of research* is the methods, algorithms, and hardware solutions used to ensure dynamic stability, including model-based control, reinforcement learning, sim-to-real transfer, state estimation, and hardware innovations.

*The purpose of the research* is to conduct a comprehensive analytical review of contemporary approaches to dynamic stability in legged robots, identify key challenges, and outline promising directions for future research.

**Methods.** To achieve this purpose, the following main *research objectives are identified*:

- analyse key scientific works on dynamic stability in legged robots, which enables a clear statement of assumptions, performance metrics, and reported disturbance-rejection behaviors;
- review commercial and open-source/academic platforms and their stability approaches, which enable mapping controller and estimation choices to typical operating scenarios;
- systematise model-based, reinforcement learning, and hybrid control strategies, which will state when each is used and what limits each approach has;
- examine sim-to-real transfer techniques, sensor fusion methods, and hardware design innovations, and note what concretely improved robustness on real robots;
- determine the scientific novelty and practical significance of the results and propose future research directions, which will specify clear next steps to pursue.

**Results and their Discussions.** This section provides an overview of leading projects and initiatives in legged robotics, divided into commercial and open-source/research efforts. Many platforms have pushed the boundaries of locomotion, dynamic stability, and real-world deployment – from logistics to entertainment and exploration. Comparison results are presented in the Table.

***Commercial Platforms***

- **1X Technologies**: Specializes in humanoid robotics with robots like *EVE* and *NEO*. [See Appendix 6.1.]
- **Agility Robotics**: Specializes in bipedal robots designed for logistics with their flagship robot *Digit*.
- **ANYbotics**: Provides autonomous solutions for industrial inspection with *ANYmal*.
- **Apptronik**: Focuses on advanced humanoid robots like *Apollo* and exoskeletons for enhanced safety and performance.
- **Boston Dynamics**: Known for dynamic stability in robots such as *Atlas*, *Spot*, and *Stretch*.
- **Clone**: Known for an android such as *Protoclone*.
- **Deep Robotics**: Specializes in high-performance quadrupeds such as *X30* and is prototyping a humanoid platform.
- **Engineered Arts**: Develops lifelike humanoid robots for interaction and entertainment (e.g., *Ameca*, *Mesmer*).
- **Figure AI**: Develops humanoid robots for industrial tasks, including *Figure 01* and *Figure 02*.
- **Fourier Intelligence**: Focuses on rehabilitation and humanoid robotics (e.g., *Fourier GR-1*, *GR-2*).
- **GITAI**: Develops space exploration robots like *GITAI-G1*.
- **Honda Robotics**: Pioneered advanced humanoid robotics with systems such as *ASIMO*.
- **Kawada Industries and AIST**: Produce humanoid and industrial robots (e.g., *HRP-2*, *HRP-4*, *Nextage*) emphasizing mobility, collaboration, and safety.
- **Kepler Robotics**: Develops the *Forerunner* K-series general-purpose humanoids for industrial automation and logistics.
- **PAL Robotics**: Develops humanoid and mobile robots such as *REEM-C*, *TALOS*, *TIAGo Pro*, and *ARI*.
- **Sanctuary AI**: Works on general-purpose humanoid robots and control systems (e.g., *Phoenix*, *Carbon AI Control System*).
- **SoftBank Robotics**: Renowned for service and social robots like *Pepper* and *NAO*.
- **Tesla**: Recently introduced the humanoid robot *Optimus* for general-purpose tasks.
- **UBTECH Robotics**: A leader in humanoid robot development (e.g., *Walker S Series*, *Walker X*).
- **Unitree Robotics**: Offers quadruped and humanoid robots with an emphasis on performance and AI integration.
- **Xiaomi Robotics Lab**: Developed platforms like *CyberDog* and *CyberDog 2* aimed at both consumer and research applications.
- **XPENG Robotics**: Develops advanced bionic robots for industrial and consumer use (e.g., *Robot Unicorn*, *PX5*).
- **Further reading:** A handy online robotics portal covers all of these platforms at a glance [3].

***Open and Academic***

- **Duke University**: Developed the *Duke Humanoid* for energy-efficient bipedal walking research.
- **Indian Institute of Science (IISc), Bengaluru**: Developed the cost-effective biped *Stoch BiRo* for uneven terrains.

- **K-Scale Labs**: Focuses on open-source humanoid robotics, such as *K-Bot 0.1*.
- **KAIST**: Developed humanoid robots like *Hubo*, *Albert Hubo*, *DRC-Hubo+*, and *PIBOT* to advance mobility, disaster response, and AI-driven piloting.
- **MIT**: Built an advanced humanoid robot capable of acrobatic behaviors.
- **NASA**: Pioneered humanoid robotics for space with *Robonaut 2* and *Valkyrie*, aiding astronaut support and extraterrestrial missions.
- **AIST**: Developed the HRP series of humanoid robots.
- **NAVER LABS**: Focuses on robotics research and developed *AMBIDEX*.
- **Noetix Robotics**: Created the *NING Humanoid* for advanced locomotion and athletic tasks.
- **PND Robotics**: Specializes in modular humanoid robots such as *Adam*.
- **Rethink Robotics**: Introduced collaborative robots such as *Baxter*.
- **RoMeLa (UCLA)**: Developed innovative robots like *CHARLI*, *DARwIn-OP*, and *THOR-RD* for bipedal locomotion, disaster response, and dynamic tasks.
- **DFKI Robotics Innovation Center**: Developed the *RH5 Humanoid Robot* for high-performance dynamic tasks.
- **Universität Bonn**: Created the *igus® Humanoid Open Platform* and *NimbRo-OP2(X)* for research and competitions.
- **USC**: Developed robots focused on autonomous locomotion and human-robot interaction.
- **University of Tehran**: Created the *Surena Humanoid Robot* series for advanced research.
- **University of Wisconsin–Madison**: The UW WELL Lab advances legged robotics through projects such as *STRIDE*.
- **Virginia Tech**: Developed *SAFFiR*, a humanoid robot for firefighting on naval ships with thermal resilience and autonomous capabilities.
- **Waseda University**: Pioneered humanoid robotics with robots like *Wabot 1*, *Wabot 2*, and *Kobian*.
- **Willow Garage**: Developed the research platform *PR2*.

**Table 1**

Robots comparison chart (sorted by year)

| Year | Organization | Name | Category | Type | Status | Height (m) | Mass (kg) | DoF | Payload (kg) | Runtime (min) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1970 | WasedaUniversity | WL-3 | Research | Biped | Retired | 1 | 60 | 6 | N/A | N/A |
| 1973 | WasedaUniversity | WABOT-1 | Research | Humanoid | Retired | 1.9 | 130 | 50 | 2 | N/A |
| 1984 | WasedaUniversity | WABOT-2 | Research | Humanoid | Retired | 1.65 | 90 | 50 | 1 | N/A |
| 1986 | Honda | E0 | Research | Biped | Retired | 1 | 50 | 6 | N/A | N/A |
| 1989 | Honda | E2 | Research | Biped | Retired | 1.2 | 70 | 6 | N/A | N/A |
| 1993 | Sony | SDR-3X(earlydev) | Research | Humanoid | Prototype | 0.5 | 6 | 20 | 0.2 | 30 |
| 1999 | Sony | AIBOERS-110 | Consumer | Quadruped(petrobot) | Discontinued | 0.27 | 1.4 | 18 | 0 | 120 |
| 2000 | HondaRobotics | ASIMO | Research | Humanoid | Discontinued | 1.2 | 52 | 34 | 5 | 40 |
| 2001 | BostonDynamics | RHex | Commercial | Hexapod | Discontinued | 0.15 | 13 | 6 | 2 | 360 |
| 2003 | KawadaIndustries | HRP-2 | Research | Humanoid | Retired | 1.54 | 58 | 30 | 2 | 30 |
| 2003 | KokoroCo. /ATR | Actroid | Research | Humanoid | Active | 1.65 | 40 | 42 | 0 | N/A |
| 2004 | Sony | QRIO | Research | Humanoid | Discontinued | 0.6 | 7.3 | 38 | 0 | 60 |
| 2005 | KAIST | Hubo | Research | Humanoid | Active | 1.21 | 43 | 41 | 2 | 40 |
| 2005 | HansonRobotics | K-Bot | Research | Humanoid | Retired | 0.55 | 3 | 15 | 0 | N/A |

| Year | Organization | Name | Category | Type | Status | Height (m) | Mass (kg) | DoF | Payload (kg) | Runtime (min) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2005 | Fujitsu | HOAP-3 | Research | Humanoid | Discontinued | 0.6 | 8.8 | 28 | 0.5 | 20 |
| 2005 | HansonRobotics | AlbertHUBO | Research | Humanoid | Prototype | 1.37 | 57 | 66 | 2 | N/A |
| 2010 | KawadaIndustries | HRP-4 | Research | Humanoid | Retired | 1.51 | 39 | 34 | 0.5 | 30 |
| 2010 | RoMeLa(UCLA) | CHARLI | Research | Humanoid | Active | 1.45 | 12 | 20 | 0.5 | 30 |
| 2010 | WillowGarage | PR2 | OpenSource | Humanoid | Inactive | 1.65 | 230 | 32 | 1.8 | 120 |
| 2011 | NASA | Robonaut2(R2) | Research | Humanoid(upperbody) | Active | 1.0(approx) | 50 | 42 | 9 | N/A |
| 2011 | RoMeLa(UCLA) | DARwIn-OP | Research | Humanoid | Active | 0.57 | 2.9 | 20 | 0.2 | 30 |
| 2011 | MIT | Cheetah(Gen1) | Research | Quadruped | Prototype | 0.6 | 15 | 12 | 5 | 15 |
| 2013 | BostonDynamics | Atlas | Research | Humanoid | Retired | 1.88 | 150 | 28 | 11 | N/A |
| 2013 | BostonDynamics | BigDog | Commercial | Quadruped | Discontinued | 0.76 | 110 | 16 | 40 | 30 |
| 2013 | NASA | Valkyrie(R5) | Research | Humanoid | Active | 1.9 | 125 | 44 | 20 | N/A |
| 2013 | RoMeLa(UCLA) | THOR | Research | Humanoid | Prototype | 1.5 | 45 | 28 | 5 | N/A |
| 2014 | BostonDynamics | WildCat | Commercial | Quadruped | Discontinued | 0.75 | 150 | 12 | 20 | 30 |
| 2014 | SoftBankRobotics | Pepper | Commercial | Humanoid | Active | 1.2 | 28 | 20 | 1 | 600 |
| 2014 | VirginiaTech | SAFFiR | Research | Humanoid | Prototype | 1.78 | 65 | 30 | 5 | N/A |
| 2015 | AgilityRobotics | Digit | Commercial | Bipedal | Active | 1.75 | 65 | 16 | 16 | 90 |
| 2015 | BostonDynamics | LS3 | Commercial | Quadruped | Discontinued | 0.91 | 508 | 12 | 180 | 1440 |
| 2015 | UniversityofTehran | SurenaIII | Research | Humanoid | Active | 1.9 | 98 | 31 | 7 | N/A |
| 2016 | ANYbotics | ANYmal | Commercial | Quadruped | Active | 0.69 | 31 | 12 | 10 | 120 |
| 2016 | RethinkRobotics | Baxter | OpenSource | Collaborative | Discontinued | 1.6 | 75 | 15 | 2.3 | AC |
| 2016 | BostonDynamics | Handle | Commercial | Wheeled | Discontinued | 2 | 105 | 10 | 45 | 60 |
| 2016 | NoetixRobotics | NINGHumanoid | OpenSource | Humanoid | Active | 1.75 | 60 | 30 | 5 | 90 |
| 2016 | XPENGRobotics | PX5 | Commercial | Humanoid | Active | 1.5 | 40 | 20 | 3 | 60 |
| 2016 | XPENGRobotics | RobotUnicorn | Commercial | Quadruped | Prototype | 0.5 | 30 | 12 | 10 | 120 |
| 2016 | PuduRobotics | PuduBot | Commercial | Service | Active | 1 | 40 | 3 | 13 | 480 |
| 2016 | HansonRobotics | Sophia | Commercial | Humanoid | Active | 1.7 | 48 | 62 | 0 | N/A |
| 2017 | FourierIntelligence | GR-1 | Commercial | Humanoid | Active | 1.65 | 55 | 40 | 5 | 120 |
| 2017 | PALRobotics | ARI | Commercial | Humanoid | Active | 1.65 | 40 | 15 | 2 | 480 |
| 2017 | PALRobotics | REEM-C | Commercial | Humanoid | Active | 1.65 | 80 | 44 | 10 | 120 |
| 2017 | BostonDynamics | Spot | Commercial | Quadruped | Active | 0.84 | 31.7 | 12 | 14 | 90 |

| Year | Organization | Name | Category | Type | Status | Height (m) | Mass (kg) | DoF | Payload (kg) | Runtime (min) |
|---|---|---|---|---|---|---|---|---|---|---|
| 2017 | BostonDynamics | SpotClassic | Commercial | Quadruped | Discontinued | 0.84 | 72 | 12 | 23 | 45 |
| 2017 | PALRobotics | TALOS | Commercial | Humanoid | Active | 1.75 | 95 | 32 | 6 | 90 |
| 2017 | PALRobotics | TIAGoPro | Commercial | Manipulator | Active | 1.45 | 70 | 10 | 3 | 480 |
| 2017 | DeepRobotics | Jueying | Commercial | Quadruped | Active | 0.5 | 20 | 12 | 5 | 120 |
| 2017 | UnitreeRobotics | Laikago | Commercial | Quadruped | Active | 0.6 | 22 | 12 | 5 | 150 |
| 2020 | UniversitätBonn | igusHumanoid | OpenSource | Humanoid | Active | 1 | 8 | 20 | 2 | 30 |
| 2020 | XiaomiRoboticsLab | CyberDog | Commercial | Quadruped | Active | 0.4 | 14 | 12 | 3 | 160 |
| 2020 | XiaomiRoboticsLab | CyberDog2 | Commercial | Quadruped | Active | 0.36 | 8.5 | 13 | 3 | 60 |
| 2020 | UniversityofWisconsin–Madison | STRIDE | Research | Quadruped | Active | 0.5 | 13 | 12 | 5 | 30 |
| 2020 | MIT | MiniCheetah | Research | Quadruped | Active | 0.3 | 9 | 12 | 2 | 20 |
| 2021 | IndianInstituteofScience | StochBiRo | OpenSource | Bipedal | Active | 0.45 | 4 | 6 | 2 | 30 |
| 2021 | 1XTechnologies | EVE | Commercial | Humanoid | Active | 1.7 | 60 | 28 | 3 | 60 |
| 2021 | 1XTechnologies | NEO | Commercial | Humanoid | Prototype | 1.7 | 55 | 25 | 3 | 60 |
| 2021 | GITAI | GITAI-G1 | Commercial | SpaceHumanoid | Active | 1.2 | 45 | 25 | 2 | N/A |
| 2021 | UnitreeRobotics | AlienGo | Commercial | Quadruped | Active | 0.5 | 23.5 | 12 | 5 | 120 |
| 2021 | UnitreeRobotics | B1 | Commercial | Quadruped | Prototype | 0.7 | 50 | 12 | 40 | 240 |
| 2021 | UnitreeRobotics | Go1 | Commercial | Quadruped | Active | 0.3 | 12 | 12 | 3 | 120 |
| 2021 | UnitreeRobotics | H1 | Commercial | Humanoid | Prototype | 1.7 | 47 | 20 | 5 | 60 |
| 2021 | EngineAIRobotics | EngineAI-X | Commercial | Humanoid | Prototype | 1.65 | 60 | 28 | 5 | 60 |
| 2021 | PNDRobotics | Adam | OpenSource | Humanoid | Active | 1.7 | 55 | 25 | 5 | 60 |
| 2021 | GhostRobotics | Vision60 | Commercial | Quadruped | Active | 0.6 | 32 | 12 | 10 | 180 |
| 2021 | BostonDynamics | Stretch | Commercial | Manipulator | Active | 2 | 500 | 7 | 23 | 480 |
| 2022 | Tesla | OptimusG1 | Commercial | Humanoid | Prototype | 1.73 | 57 | 28 | 20 | 60 |
| 2022 | GalbotRobotics | GalbotExplorer | Commercial | Quadruped | Prototype | 0.6 | 25 | 12 | 8 | 90 |
| 2022 | MenteeRobotics | MenteeOne | Commercial | Humanoid | Prototype | 1.6 | 50 | 25 | 5 | 60 |
| 2022 | XiaomiRoboticsLab | CyberOne | Commercial | Humanoid | Prototype | 1.77 | 52 | 21 | 1.5 | 45 |
| 2023 | Tesla | OptimusG2 | Commercial | Humanoid | Prototype | 1.73 | 47 | 28 | 20 | 60 |
| 2023 | FigureAI | Figure01 | Commercial | Humanoid | Prototype | 1.68 | 60 | 40 | 20 | 300 |
| 2023 | Apptronik | Apollo | Commercial | Humanoid | Prototype | 1.73 | 73 | 30 | 25 | 240 |
| 2023 | LimXDynamics | LimXInfinity | Commercial | Humanoid | Active | 1.75 | 85 | 52 | 10 | 480 |
| 2023 | SUPCON | SUPCONIris | Commercial | Humanoid | Prototype | 1.7 | 65 | 30 | 8 | 90 |
| 2024 | KeplerRobotics | ForerunnerK2 | Commercial | Humanoid | Active | 1.78 | 85 | 52 | 15 | 480 |

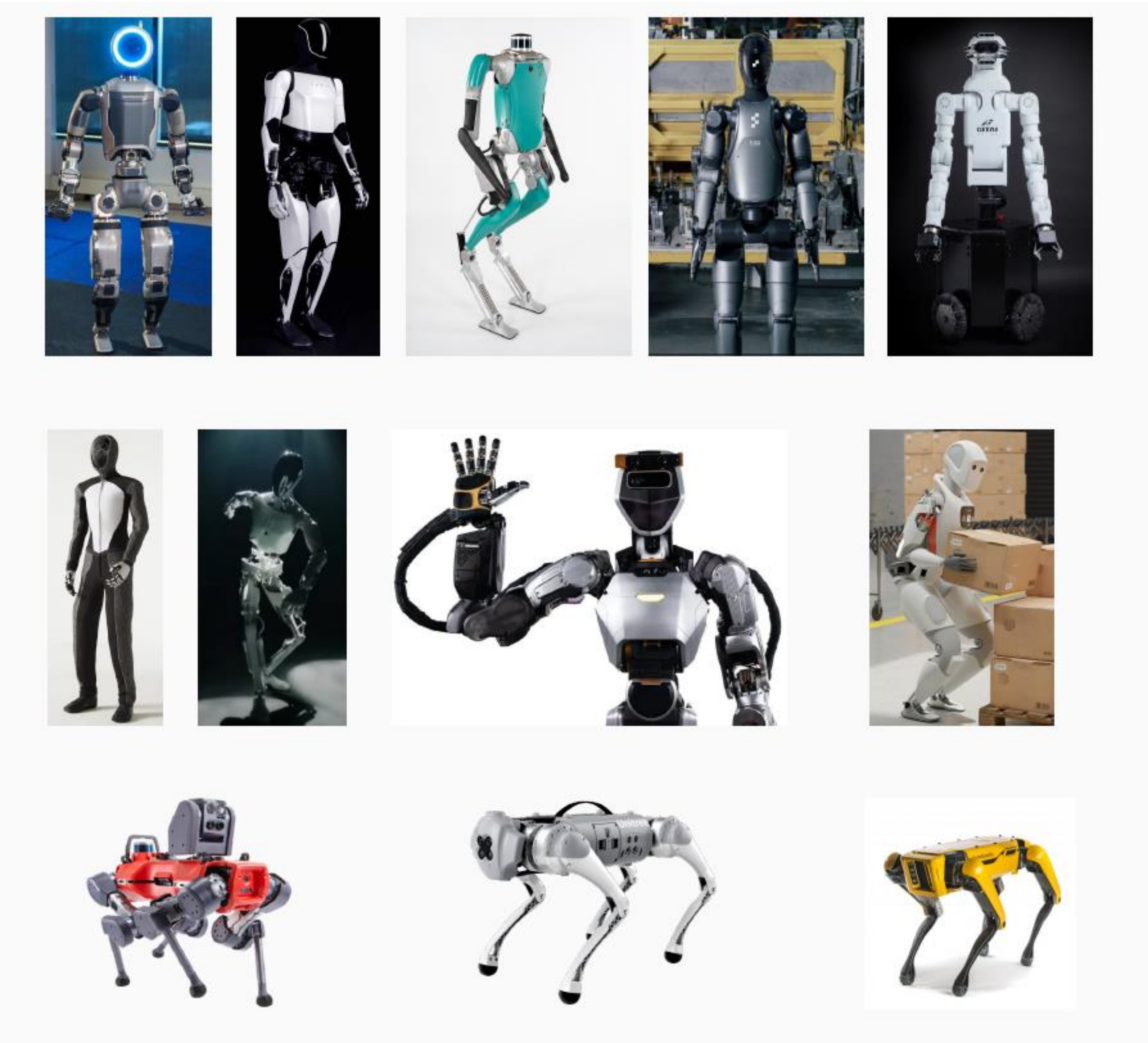

**Figure**. Atlas (Boston Dynamics), Optimus (Tesla), Digit (Agility Robotics), Figure 02 (Figure AI), G1 (Gitai), NEO (1X Technologies), Iron (XPENG Robotics), Phoenix (Sanctuary AI), Apollo (Apptronik), ANYmal (ANYbotics), Go1 (Unitree Robotics), Spot (Boston Dynamics)

**Analysis of recent research and publications**. Hezog et al. implement hierarchical inverse dynamics on a torque-controlled humanoid and show reliable momentum/contact tracking during walking [4]. Escande et al. propose a fast hierarchical QP and enable online whole-body motion at control rates suitable for balance [5]. Dai, Valenzuela, and Tedrake combine centroidal dynamics with full kinematics to plan feasible whole-body motions for locomotion [6]. Kajita and Tani introduce the LIPM and demonstrate stable bipedal walking on uneven terrain with a simplified template [7]. Haddadin surveys safety principles for contact-rich interaction and defines constraints for manipulation controllers [1]. Calinon et al. use imitation learning to reproduce human-like arm motions that generalize to new task conditions [8]. Ott et al. present Cartesian impedance control for flexible-joint robots and show compliant, safe interaction via stiffness and damping regulation [9]. Levine et al. train end-to-end visuomotor policies from real data and achieve closed-loop arm control directly from pixels [10]. Kuindersma et al. integrate optimization-based planning, estimation, and control on Atlas and demonstrate coordinated multi-contact behaviors [11]. Yang et al. improve data efficiency for reinforcement learning in legged systems and reduce the samples needed for locomotion skills [12]. Fallon et al. couple online affordance-based perception with whole-body planning to select feasible actions in real time [13]. Atkeson and Stephens analyze random sampling for dynamic programming and provide tools for coordinating high-dimensional control [14].

The surveyed works cover model-based control (LIPM, centroidal dynamics, hierarchical inverse dynamics, MPC), learning-based controllers (PPO, SAC, DDPG, TD3, transformer policies), sim-to-real techniques (domain/dynamics randomization, curricula, RMA), state estimation (EKF/IEKF, multiple-model, invariant neural-augmented filters), hardware choices (compliant/series-elastic actuation, thruster assistance), and safety elements (constrained policy optimization, backup controllers). However, results are reported under heterogeneous assumptions, metrics, and contact conditions, and are spread across commercial, open-source, and academic platforms without a unified view. This study contributes a structured synthesis and a consolidated platform comparison, linking published methods to typical scenarios (flat ground, rough terrain, payload changes, and loco-manipulation) and summarizing open challenges highlighted in the sources. The need for such an aligned, evidence-based overview motivates the present work.

**Classical and Hybrid Control.** Traditional control strategies rely on model-based techniques that use the robot's dynamics to keep it stable [15]. The Linear Inverted Pendulum Model (LIPM) is widely used to generate walking patterns in bipedal robots by simplifying their dynamics [7], [108]. Extensions that include centroidal dynamics and full kinematics support whole-body motion during locomotion and manipulation [6]. Hierarchical inverse-dynamics and quadratic programming formulations help improve stability while juggling multiple tasks [4], [5], [16].

A recent MuJoCo benchmark compared a Unitree Go1 controlled by preview-horizon MPC with the same robot using PPO-based RL under identical external pushes. Within that setup, MPC handled larger disturbances thanks to explicit optimisation, whereas RL achieved a lower cost of transport on flat ground. These results are indicative of that configuration (controller tuning, horizons, reward/cost weights, and simulator fidelity) and do not establish a general ranking; rather, they suggest that a hybrid MPC–RL pipeline could capture the benefits of both approaches [17].

Safe and compliant interaction with the environment is also important. Impedance control lets robots adjust joint stiffness and damping, helping them absorb impacts and stay stable under contact forces [9]. Collision-detection and safe-response methods further contribute to both safety and stability [1].

**Learning-Based Methods.** Reinforcement learning (RL) lets robots refine control policies through trial and error [18]. Its roots include early work on temporal-difference learning and natural policy gradients [19], [20], [21], [22]. Modern model-free algorithms – DQN [23], [24], DRQN [25], PPO [26], DDPG [27], and SAC [28] – have trained quadrupeds to run, obey velocity commands, and stand up after falls [29]. A vision-based PPO policy later let the same robot traverse rocks and steps at 1 m/s using only onboard sensors [30]. Large MuJoCo studies mapped sample-efficiency and stability trade-offs across these algorithms [31], [32]. Data-efficient tools such as Hindsight Experience Replay [33], imitation learning (DeepMimic) [34], and hierarchical policies [12], [35] shorten training time.

Recent controllers now match full-size bipeds. A transformer policy trained in simulation let Agility Digit walk outdoors and resist pushes without further tuning [36]. Imitation-Relaxation Reinforcement Learning (IRRL) reached 5 m/s on a Mini-Cheetah while keeping gaits stable [37]. For value methods, Double DQN [38], Dueling networks [39], Prioritized Replay [40], and Rainbow [41] cut bias and speed learning. A3C [42] and Evolution Strategies [43] scale training; TRPO [44], K-FAC TRPO [45], and PPO keep on-policy updates stable. TD3 mitigates critic over-estimation [46]; PCL unifies actor–critic and Q-learning under one consistency loss [47]. Conservative Q-Learning tackles over-estimation in offline settings [48]. MAVIPER adds interpretability with decision-tree policies [49]. Beyond control, AlphaZero proved RL's reach in board games [50].

Imitation learning avoids hand-crafted rewards. GAIL frames imitation as a GAN game [51]. Even a simple speed reward can yield rich behaviours when combined with procedural terrains [52].

Whole-body loco-manipulation has also advanced. Digit lifted and placed boxes after simulation training [53]; a quadruped-arm system learned smooth coordination [54]; WoCoCo used phase-specific networks for humanoid motion [55]; PolyTask distilled several specialists into one multi-task policy [56].

Meta-learning cuts adaptation time. MAML [57] and $RL^{2}$ [58] fine-tune quickly. Classic tools such as LQR-trees [59], [60] and distributional RL [61], [62] refine control. Hybrid controllers pair parametric gait generators with residual networks [63] or planners with RL execution [64]. Conditioned policies switch gaits on the fly [65]. Curriculum plus meta-RL extend speed ranges: curricular HER quadrupeds ran 3.45 m/s outside [66], and MetaLoco gave zero-shot walking across designs [67]. Rapid Motor Adaptation copes with sudden terrain changes [68].

Effective sim-to-real transfer is still vital. Progressive Nets reduce forgetting while bridging the reality gap [69], [70]. Policy distillation and Actor-Mimic compress multiple gaits [71], [72]. Option-Critic and Meta-Learning Shared Hierarchies reuse low-level skills [73], [74]. Latent-space MPC [75] and curriculum-based HER [76] further boost efficiency.

Overall, learning-based methods now rival classical controllers in many tasks and offer fast adaptation. Progress in data efficiency, safety, and transfer should make them even more practical for real-legged robots.

**Sim-to-Real, Hardware, and Perception.** Policies that work in simulation often fail on hardware because simulators miss many subtle effects. Domain and dynamics randomisation make controllers see a wider range of conditions, which improves robustness [77], [78]. With heavy randomisation alone, Tan transferred a quadruped policy from MuJoCo to the real robot on the first try [79].

Rapid Motor Adaptation (RMA) addresses changing terrain by pairing a base policy with an online adaptation module [80]. Trained only in simulation, the same quadruped walks on grass, stairs, pebbles, and sand with no manual tuning. The idea relates to meta-RL adaptation modules [81]. Dao [82] randomised payload mass so a biped kept its footing while pulling carts or carrying liquids, and Duan [64] learned a transition model that respects footstep plans on Cassie. Model-based meta-policy optimisation speeds such adaptation by learning a simulator-agnostic critic [83].

Curricula and stochastic contact modelling also help. A box-manipulation curriculum let Digit lift and place boxes without extra system identification [53]. NVIDIA’s GR00T N1 takes a broader step: it combines a vision–language transformer with a diffusion action head and executes language-conditioned tasks on a GR-1 humanoid after purely simulated training [84].

Hardware choices can boost stability, too. Caltech’s LEO mixes legs with small thrusters for extra balance [85]. Series-elastic and compliant actuators absorb shocks on rough ground [86], and elastic hips with passive knees and ankles store energy in simulation tests [87]. A broad survey of recent humanoid hardware appears in [88]. Open designs such as Solo [89] and the series–parallel RH5 layout [90] show how lightweight drives and compliant linkages improve bandwidth and payload capacity.

Accurate state estimation is another pillar. Combining leg kinematics with inertial data in an EKF offers real-time estimates [91], [92], while multiple-model filters handle changing contact modes [93]. Hybrid filters such as the Invariant Neural-Augmented KF (InNKF) limit drift after slips [94], [95], and low-rate camera or LiDAR updates, as in Pronto, bound long-term error [96]. Large-scale real-world data can help as well: over 800,000 grasp attempts let a vision network learn closed-loop grasping without precise calibration [97].

Finally, safety remains vital. Constrained Policy Optimisation [98] and intrinsic-motivation exploration [99], [100] reduce risky actions, yet learned policies still lack formal guarantees [101]. One remedy adds a backup controller that takes over near constraints [102]. Others learn Lyapunov or control-barrier functions alongside the main policy [101]. These hybrids lower the fall rate in hardware tests while keeping task performance high.

Together, advances in transfer, hardware, perception, and safety continue to narrow the gap between simulation and reliable real-world deployment.

**Synthesis.** Recent studies – such as [36], [37], [63], [64], [66], [67], [68], [80], [82], [102] – show steady progress toward reliable, agile legged robots. Learning-based controllers now deliver fast, adaptable gaits on rough terrain, while sim-to-real techniques narrow the modelling gap. Hardware innovations (e.g. thrusters, soft or series-elastic actuators) extend the range of stable behaviours. Improved state-estimation pipelines fuse contact cues with neural corrections to handle slip. Safety layers – backup policies or control-barrier functions – cut fall rates. Hybrid designs combine model-based planning with data-driven residuals, and meta-learning speeds online adaptation. Together, these strands bring legged machines closer to animal-level stability, letting them cope with unpredictable disturbances and new tasks that once defeated purely model-based methods.

For broader context, Tong reviews recent humanoid hardware and control trends [103]; Ha surveys learning-based legged locomotion and open problems [104]; and Gu outline a 2025 roadmap for humanoid locomotion and manipulation that integrates planning, control, and learning [105].

**Future Directions.** Robotic dynamic stability is an evolving field with several areas that require further exploration:

- **Hybrid Control Architectures:** Integrate reinforcement learning with model-based control [106] to combine adaptability with stability.
- **Adaptability and Online Tuning:** Techniques such as the Multiplicity of Behavior (MoB) framework [107] highlight the need for real-time tuning.
- **Extreme Dynamic Stability:** Research such as line walking with point feet [108] opens new avenues for maintaining stability on minimal supports.
- **Multi-Legged Locomotion:** Extending controllers to robots with many legs [109] can improve both speed and dynamic stability.
- **Sim-to-Real Transfer:** Enhancing simulation fidelity and developing robust transfer methods remain essential for bridging the gap between simulation and reality.

Collaboration across academia, industry, and policy is essential to address these challenges and drive the next phase of legged robotics.

**Conclusions.** We met our goal: reviewed approaches to dynamic stability in legged robots, identified the main challenges, and outlined next steps.

Based on the results of the research, the following main conclusions can be drawn:

-systematized model-based, learning-based, and hybrid control to show how each supports gait, balance, and whole-body motion;

-analyzed published results on walking, manipulation, and coordinated behaviors, relying only on the cited sources;

-clarified sim-to-real tools — domain and dynamics randomization, curricula, distillation, progressive nets, and rapid motor adaptation — and what they improve for hardware transfer;

-showed that estimation pipelines (leg-IMU EKF/IEKF, multiple-model/contact-aided, invariant neural-augmented filters) are central for cutting drift and handling contact changes;

-highlighted hardware choices — compliant or series-elastic actuation and thruster assistance — that widen the range of recoverable disturbances;

-presented a single comparison of commercial, open, and academic platforms by their reported stability practices and specifications;

Taken together, the literature points to hybrid control plus strong estimation as key for stability under intermittent contacts, while sim-to-real fidelity and safety layers remain practical bottlenecks. Near-term work should tighten model–RL integration, enable online adaptation, explore point-foot balancing, and extend methods to multi-leg systems. The practical value is a scenario-based view that helps pick controllers, transfer methods, estimators, and safety layers for real robots.

## References:


1. Haddadin, S. (2013). *Towards safe robots: Approaching Asimov's first law*. Springer

2. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv*. https://arxiv.org/abs/1606.06565

3. Robots Guide. (n.d.). *Robots: Your guide to the world of robotics*. https://robotsguide.com

4. Herzog, A., Rotella, N., Mason, S., Grimminger, F., Schaal, S., & Righetti, L. (2015). Momentum control with hierarchical inverse dynamics on a torque-controlled humanoid. *Autonomous Robots, 40*(3), 473–491. https://doi.org/10.1007/s10514-015-9476-6

5. Escande, A., Mansard, N., & Wieber, P.-B. (2014). Hierarchical quadratic programming: Fast online humanoid-robot motion generation. *The International Journal of Robotics Research, 33*, 1006–1028. doi.org/10.1177/0278364914521306

6. Dai, H., Valenzuela, A., & Tedrake, R. (2014). Whole-body motion planning with centroidal dynamics and full kinematics. In *2014 IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (pp. 295–302). IEEE. https://doi.org/10.1109/HUMANOIDS.2014.7041375

7. Kajita, S., & Tani, K. (1991). Study of dynamic biped locomotion on rugged terrain: Derivation and application of the linear inverted pendulum mode. In *1991 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1405–1411). IEEE. https://doi.org/10.1109/ROBOT.1991.131811

8. Calinon, S., D'halluin, F., Sauser, E. L., Caldwell, D. G., & Billard, A. G. (2010). Learning and reproduction of gestures by imitation. *IEEE Robotics & Automation Magazine, 17*(2), 44–54. https://doi.org/10.1109/MRA.2010.936947

9. Ott, C., Albu-Schaffer, A., Kugi, A., Stamigioli, S., & Hirzinger, G. (2004). A passivity-based Cartesian impedance controller for flexible-joint robots—Part I: Torque feedback and gravity compensation. In *2004 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2659–2665). IEEE. https://doi.org/10.1109/ROBOT.2004.1307462.

10. Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. *arXiv preprint arXiv:1504.00702*. https://arxiv.org/abs/1504.00702

11. Kuindersma, S., Deits, R., Fallon, M., Valenzuela, A., Dai, H., Permenter, F., Koolen, T., Marion, P., & Tedrake, R. (2015). Optimization-based locomotion planning, estimation, and control design for the Atlas humanoid robot. *Autonomous Robots, 40*. https://doi.org/10.1007/s10514-015-9479-3

12. Yang, Y., Caluwaerts, K., Iscen, A., Zhang, T., Tan, J., & Sindhwani, V. (2019). Data-efficient reinforcement learning for legged robots. *arXiv preprint arXiv:1907.03613*. https://arxiv.org/abs/1907.03613

13. Fallon, M., Kuindersma, S., Karumanchi, S., Antone, M., Schneider, T., Dai, H., D'Arpino, C., Deits, R., DiCicco, M., Fourie, D., Koolen, T., Marion, P., Posa, M., Valenzuela, A., Yu, K.-T., Shah, J., Iagnemma, K., Tedrake, R., & Teller, S. (2014). An architecture for online affordance-based perception and whole-body planning. *Journal of Field Robotics, 32*. https://doi.org/10.1002/rob.21546

14. Atkeson, C. G., & Stephens, B. J. (2008). Random sampling of states in dynamic programming. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 38*(4), 924–929. https://doi.org/10.1109/TSMCB.2008.926610.

15. Deisenroth, M. P., Neumann, G., & Peters, J. (2013). *A survey on policy search for robotics*. https://doi.org/10.1561/2300000021

16. Peters, J., & Schaal, S. (2008). Natural actor-critic. *Neurocomputing, 71*, 1180–1190. https://doi.org/10.1016/j.neucom.2007.11.026

17. Akki, S., & Chen, T. (2025). Benchmarking model predictive control and reinforcement learning based control for legged robot locomotion in MuJoCo simulation. *arXiv preprint arXiv:2501.16590*. https://arxiv.org/abs/2501.16590

18. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. A Bradford Book.

19. Barto, A., Bradtke, S., & Singh, S. (1993). Learning to act using real-time dynamic programming. *Artificial Intelligence, 72*, 81–138. https://doi.org/10.1016/0004-3702(94)00011-O

20. Ng, A. Y., & Russell, S. J. (2000). Algorithms for inverse reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning* (pp. 663–670). San Francisco, CA: Morgan Kaufmann Publishers.

21. Kakade, S. M. (2001). A natural policy gradient. In T. Dietterich, S. Becker, & Z. Ghahramani (Eds.), *Advances in Neural Information Processing Systems* (Vol. 14). Kakade, S., & Langford, J. (2002). Approximately optimal approximate reinforcement learning. In *Proceedings of the Nineteenth International Conference on Machine Learning* (ICML).

22. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing Atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*. https://arxiv.org/abs/1312.5602

23. Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. *Nature*. https://doi.org/10.1038/nature14236

24. Hausknecht, M., & Stone, P. (2017). Deep recurrent Q-learning for partially observable MDPs. *arXiv preprint arXiv:1507.06527*. https://arxiv.org/abs/1507.06527

25. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*. https://arxiv.org/abs/1707.06347

26. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., & Wierstra, D. (2019). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*. https://arxiv.org/abs/1509.02971

27. Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *arXiv preprint arXiv:1801.01290*. https://arxiv.org/abs/1801.01290

28. Hwangbo, J., Lee, J., Dosovitskiy, A., Bellicoso, D., Tsounis, V., Koltun, V., & Hutter, M. (2019). Learning agile and dynamic motor skills for legged robots. *Science Robotics, 4*(26). https://doi.org/10.1126/scirobotics.aau5872

29. Lee, J., Hwangbo, J., Wellhausen, L., Koltun, V., & Hutter, M. (2020). Learning quadrupedal locomotion over challenging terrain. *Science Robotics, 5*(47). doi.org/10.1126/scirobotics.abc5986

30. Duan, Y., Chen, X., Houthooft, R., Schulman, J., & Abbeel, P. (2016). Benchmarking deep reinforcement learning for continuous control. *arXiv preprint arXiv:1604.06778*. https://arxiv.org/abs/1604.06778

31. Kober, J., Bagnell, J., & Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research, 32*, 1238–1274. doi.org/10.1177/0278364913495721

32. Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, R., Welinder, P., McGrew, B., Tobin, J., Abbeel, P., & Zaremba, W. (2018). Hindsight experience replay. *arXiv preprint arXiv:1707.01495*. https://arxiv.org/abs/1707.01495

33. Peng, X. B., Abbeel, P., Levine, S., & van de Panne, M. (2018). DeepMimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions on Graphics, 37*(4), 1–14. https://doi.org/10.1145/3197517.3201311

34. Peng, X. B., Berseth, G., Yin, K., & Van De Panne, M. (2017). DeepLoco: Dynamic locomotion skills using hierarchical deep reinforcement learning. *ACM Transactions on Graphics, 36*(4), 41:1–41:13. https://doi.org/10.1145/3072959.3073602

35. Radosavovic, I., Xiao, T., Zhang, B., Darrell, T., Malik, J., & Sreenath, K. (2024). Real-world humanoid locomotion with reinforcement learning. *Science Robotics, 9*(89), eadi9579. https://doi.org/10.1126/scirobotics.adi9579

36. Jin, Y., Liu, X., Shao, Y., Wang, H., & Yang, W. (2022). High-speed quadrupedal locomotion by imitation-relaxation reinforcement learning. *Nature Machine Intelligence, 4*(12), 1198–1208. https://doi.org/10.1038/s42256-022-00576-3

37. Van Hasselt, H., Guez, A., & Silver, D. (2015). Deep reinforcement learning with double Q-learning. *arXiv preprint arXiv:1509.06461*. https://arxiv.org/abs/1509.06461

38. Wang, Z., Schaul, T., Hessel, M., Van Hasselt, H., Lanctot, M., & De Freitas, N. (2016).

Dueling network architectures for deep reinforcement learning. *arXiv preprint arXiv:1511.06581*. https://arxiv.org/abs/1511.06581

39. Schaul, T., Quan, J., Antonoglou, I., & Silver, D. (2016). Prioritized experience replay. *arXiv preprint arXiv:1511.05952*. https://arxiv.org/abs/1511.05952

40. Hessel, M., Modayil, J., Van Hasselt, H., Schaul, T., Ostrovski, G., Dabney, W., Horgan, D., Piot, B., Azar, M., & Silver, D. (2017). Rainbow: Combining improvements in deep reinforcement learning. *arXiv preprint arXiv:1710.02298*. https://arxiv.org/abs/1710.02298

41. Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T. P., Harley, T., Silver, D., & Kavukcuoglu, K. (2016). Asynchronous methods for deep reinforcement learning. *arXiv preprint arXiv:1602.01783*. https://arxiv.org/abs/1602.01783

42. Salimans, T., Ho, J., Chen, X., Sidor, S., & Sutskever, I. (2017). Evolution strategies as a scalable alternative to reinforcement learning. *arXiv preprint arXiv:1703.03864*. https://arxiv.org/abs/1703.03864

43. Schulman, J., Levine, S., Moritz, P., Jordan, M. I., & Abbeel, P. (2017). Trust region policy optimization. *arXiv preprint arXiv:1502.05477*. https://arxiv.org/abs/1502.05477

44. Wu, Y., Mansimov, E., Liao, S., Grosse, R., & Ba, J. (2017). Scalable trust-region method for deep reinforcement learning using Kronecker-factored approximation. *arXiv preprint arXiv:1708.05144*. https://arxiv.org/abs/1708.05144

45. Fujimoto, S., Van Hoof, H., & Meger, D. (2018). Addressing function approximation error in actor-critic methods. *arXiv preprint arXiv:1802.09477*. https://arxiv.org/abs/1802.09477

46. Nachum, O., Norouzi, M., Xu, K., & Schuurmans, D. (2017). Bridging the gap between value and policy based reinforcement learning. *arXiv preprint arXiv:1702.08892*. https://arxiv.org/abs/1702.08892

47. Kumar, A., Zhou, A., Tucker, G., & Levine, S. (2020). Conservative Q-learning for offline reinforcement learning. *arXiv preprint arXiv:2006.04779*. https://arxiv.org/abs/2006.04779

48. Milani, S., Zhang, Z., Topin, N., Shi, Z. R., Kamhoua, C., Papalexakis, E. E., & Fang, F. (2022). MAVIPER: Learning decision tree policies for interpretable multi-agent reinforcement learning. *arXiv preprint arXiv:2205.12449*. https://doi.org/10.48550/arXiv.2205.12449

49. Silver, D., Schrittwieser, J., Simonyan, K., et al. (2017). Mastering the game of Go without human knowledge. *Nature, 550*, 354–359. https://doi.org/10.1038/nature24270

50. Ho, J., & Ermon, S. (2016). Generative adversarial imitation learning. *arXiv preprint arXiv:1606.03476*. https://arxiv.org/abs/1606.03476

51. Heess, N., TB, D., Sriram, S., Lemmon, J., Merel, J., Wayne, G., Tassa, Y., Erez, T., Wang, Z., Eslami, S. M. A., Riedmiller, M., & Silver, D. (2017). Emergence of locomotion behaviours in rich environments. *arXiv preprint arXiv:1707.02286*. https://arxiv.org/abs/1707.02286

52. Dao, J., Duan, H., & Fern, A. (2023). Sim-to-real learning for humanoid box loco-manipulation. *arXiv preprint arXiv:2310.03191*. https://arxiv.org/abs/2310.03191

53. Fu, Z., Cheng, X., & Pathak, D. (2022). Deep whole-body control: Learning a unified policy for manipulation and locomotion. *arXiv preprint arXiv:2210.10044*. https://arxiv.org/abs/2210.10044

54. Zhang, C., Xiao, W., He, T., & Shi, G. (2024). WoCoCo: Learning whole-body humanoid control with sequential contacts. *arXiv preprint arXiv:2406.06005*. https://arxiv.org/abs/2406.06005

55. Haldar, S., & Pinto, L. (2023). PolyTask: Learning unified policies through behavior distillation. *arXiv preprint arXiv:2310.08573*. https://arxiv.org/abs/2310.08573

56. Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *arXiv preprint arXiv:1703.03400*. https://arxiv.org/abs/1703.03400

57. Duan, Y., Schulman, J., Chen, X., Bartlett, P. L., Sutskever, I., & Abbeel, P. (2016). RL^2: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint arXiv:1611.02779*. https://arxiv.org/abs/1611.02779

58. Tedrake, R., Manchester, I. R., Tobenkin, M., & Roberts, J. W. (2010). LQR-trees: Feedback motion planning via sums-of-squares verification. *The International Journal of Robotics Research, 29*(8), 1038–1052. https://doi.org/10.1177/0278364910369189

59. Berkenkamp, F., Turchetta, M., Schoellig, A. P., & Krause, A. (2017). Safe model-based reinforcement learning with stability guarantees. *arXiv preprint arXiv:1705.08551*. https://arxiv.org/abs/1705.08551

60. Bellemare, M. G., Dabney, W., & Munos, R. (2017). A distributional perspective on reinforcement learning. *arXiv preprint arXiv:1707.06887*. https://arxiv.org/abs/1707.06887

61. Dabney, W., Rowland, M., Bellemare, M. G., & Munos, R. (2017). Distributional reinforcement

learning with quantile regression. *arXiv preprint arXiv:1710.10044*.
https://arxiv.org/abs/1710.10044

62. Kasaei, M., Abreu, M., Lau, N., Pereira, A., Reis, L. P., & Li, Z. (2023). Learning hybrid locomotion skills—Learn to exploit residual actions and modulate model-based gait control. *Frontiers in Robotics and AI, 10*. https://doi.org/10.3389/frobt.2023.1004490

63. Duan, H., Malik, A., Dao, J., Saxena, A., Green, K., Siekmann, J., Fern, A., & Hurst, J. (2022). Sim-to-real learning of footstep-constrained bipedal dynamic walking. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.

64. Wei, W., Wang, Z., Xie, A., Wu, J., Xiong, R., & Zhu, Q. (2023). Learning gait-conditioned bipedal locomotion with motor adaptation. In *Proceedings of the 2023 IEEE-RAS 22nd International Conference on Humanoid Robots (Humanoids)*(pp.1–7).
//doi.org/10.1109/Humanoids57100.2023.10375167

65. Li, S., Wang, G., Pang, Y., Bai, P., Hu, S., Liu, Z., Wang, L., & Li, J. (2024). Learning agility and adaptive legged locomotion via curricular hindsight reinforcement learning. *Scientific Reports, 14*(1), 28089. https://doi.org/ 10.1038/s41598-024-79292-4

66. Zargarbashi, F., Di Giuro, F., Cheng, J., Kang, D., Sukhija, B., & Coros, S.
(2024). MetaLoco: Universal quadrupedal locomotion with meta-reinforcement learning and motion imitation. *arXiv preprint arXiv:2407.17502*. https://arxiv.org/ abs/2407.17502

67. Choi, S., Ji, G., Park, J., Kim, H., Mun, J., Lee, J. H., & Hwangbo, J. (2023). Learning quadrupedal locomotion on deformable terrain. *Science Robotics, 8*(74), eade2256. https://doi.org/10.1126/scirobotics.ade2256

68. Rusu, A. A., Rabinowitz, N. C., Desjardins, G., Soyer, H., Kirkpatrick, J., Kavukcuoglu, K., Pascanu, R., & Hadsell, R. (2016). Progressive neural networks. *arXiv preprint arXiv:1606.04671*. https://arxiv.org/abs/1606.04671

69. Rusu, A. A., Vecerik, M., Rothörl, T., Heess, N., Pascanu, R., & Hadsell, R. (2018). Sim-to-real robot learning from pixels with progressive nets. *arXiv preprint arXiv:1610.04286*. https://arxiv.org/ abs/1610.04286

70. Rusu, A. A., Gomez Colmenarejo, S., Gulcehre, C., Desjardins, G., Kirkpatrick, J., Pascanu, R., Mnih, V., Kavukcuoglu, K., & Hadsell, R. (2016). Policy distillation. *arXiv preprint arXiv:1511.06295*. arxiv.org/ abs/1511.06295

71. Parisotto, E., Ba, J. L., & Salakhutdinov, R. (2016). Actor-mimic: Deep multitask and transfer reinforcement learning. *arXiv preprint arXiv:1511.06342*. https://arxiv.org/abs/1511.06342

72. Bacon, P.-L., Harb, J., & Precup, D. (2016). The option-critic architecture. *arXiv preprint arXiv:1609.05140*.
https://arxiv.org/abs/1609.05140

73. Frans, K., Ho, J., Chen, X., Abbeel, P., & Schulman, J. (2017). Meta-learning shared hierarchies. *arXiv preprint arXiv:1710.09767*. https://arxiv.org/abs/1710.09767

74. Hafner, D., Lee, K., Fischer, I., & Abbeel, P. (2022). Deep hierarchical planning from pixels. *arXiv preprint arXiv:2206.04114*. https://arxiv.org/abs/2206.04114

75. Manela, B., & Biess, A. (2020). Curriculum learning with hindsight experience replay for sequential object manipulation tasks. *arXiv preprint arXiv:2008.09377*.
https://arxiv.org/abs/2008.09377

76. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P. (2017). Domain randomization for transferring deep neural networks from simulation to the real world. *arXiv preprint arXiv:1703.06907*.
https://arxiv.org/abs/1703.06907

77. Peng, X. B., Andrychowicz, M., Zaremba, W., & Abbeel, P. (2018). Sim-to-real transfer of robotic control with dynamics randomization. In *2018 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3803–3810). https://doi.org/10.1109/ICRA.2018.8460528

78. Tan, J., Zhang, T., Coumans, E., Iscen, A., Bai, Y., Hafner, D., Bohez, S., & Vanhoucke, V. (2018). Sim-to-real: Learning agile locomotion for quadruped robots. *arXiv preprint arXiv:1804.10332*.
https://arxiv.org/abs/1804.10332

79. Kumar, A., Fu, T., Pathak, D., & Malik, J. (2021). RMA: Rapid motor adaptation for legged robots. In *Robotics: Science and Systems (RSS)*. https://ashish-kmr.github.io/rma/

80. Nagabandi, A., Clavera, I., Liu, S., Fearing, R. S., Abbeel, P., Levine, S., & Finn, C. (2019). Learning to adapt in dynamic, real-world environments through meta-reinforcement learning. *arXiv preprint arXiv:1803.11347*. https://arxiv.org/ abs/1803.11347

81. Dao, J., Green, K., Duan, H., Fern, A., & Hurst, J. (2022). Sim-to-real learning for bipedal locomotion under unsensed dynamic loads. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1–8). https://doi.org/10.1109/ICRA46639.2022.9811783

82. Clavera, I., Rothfuss, J., Schulman, J., Fujita, Y., Asfour, T., & Abbeel, P. (2018). Model-based reinforcement learning via meta-policy

optimization. *arXiv preprint arXiv:1809.05214*. https://arxiv.org/abs/1809.05214

83. NVIDIA, Bjorck, J., Castañeda, F., Cherniadev, N., Da, X., Ding, R., & Zhu, Y. (2025). GR00T N1: An open foundation model for generalist humanoid robots. *arXiv preprint arXiv:2503.14734*. https://arxiv.org/abs/2503.14734 Caltech. (2021). LEONARDO (LEO) bipedal robot with thrusters. *Caltech News*. https://www.caltech.edu/about/ news/leonardo

84. Sherrod, V., Johnson, C. C., & Killpack, M. D. (2022). Design optimization for rough terrain traversal using a compliant, continuum-joint, quadruped robot. *Frontiers in Robotics and AI, 9*. https://doi.org/10.3389/frobt.2022.860020

85. Zhu, H., & Thomas, U. (2023). Mechanical design of a biped robot FORREST and an extended capture-point-based walking pattern generator. *Robotics, 12*(3), 82. https://doi.org/10.3390/robotics12030082

86. Ficht, G., & Behnke, S. (2021). Bipedal humanoid hardware design: A technology review. *arXiv preprint arXiv:2103.04675*. https://arxiv.org/abs/2103.04675

87. Grimminger, F., Meduri, A., Khadiv, M., Viereck, J., Wuthrich, M., Naveau, M., & Righetti, L. (2020). An open torque-controlled modular robot architecture for legged locomotion research. *IEEE Robotics and Automation Letters, 5*(2), 3650–3657. https://doi.org/10.1109/LRA.2020.2976639

88. Esser, J., Kumar, S., Peters, H., Bargsten, V., de Gea Fernandez, J., Mastalli, C., Stasse, O., & Kirchner, F. (2021). Design, analysis and control of the series-parallel hybrid RH5 humanoid robot. *arXiv preprint arXiv:2101.10591*. https://arxiv.org/abs/2101.10591

89. Bloesch, M., Hutter, M., Hoepflinger, M., Leutenegger, S., Gehring, C., Remy, C., & Siegwart, R. (2012). State estimation for legged robots: Consistent fusion of leg kinematics and IMU. In *Robotics: Science and Systems VIII*. https://doi.org/10.15607/RSS.2012.VIII.003

90. Chilian, A., Hirschmüller, H., & Görner, M. (2011). Multisensor data fusion for robust pose estimation of a six-legged walking robot. In *IEEE International Conference on Intelligent Robots and Systems* (pp. 2497–2504). https://doi.org/10.1109/IROS.2011.6048125

91. Menner, M., & Berntorp, K. (2024). Simultaneous state estimation and contact detection for legged robots by multiple-model Kalman filtering. *arXiv preprint arXiv:2404.03444*. https://arxiv.org/abs/2404.03444

92. Lee, S., Kim, H.-B., & Kim, K.-S. (2025). Legged robot state estimation using invariant neural-augmented Kalman filter with a neural compensator. *arXiv preprint arXiv:2503.00344*. https://arxiv.org/abs/2503.00344

93. Youm, D., Oh, H., Choi, S., Kim, H., & Hwangbo, J. (2024). Legged robot state estimation with invariant extended Kalman filter using neural measurement network. *arXiv preprint arXiv:2402.00366*. https://arxiv.org/abs/2402.00366

94. Camurri, M., Ramezani, M., Nobili, S., & Fallon, M. (2020). Pronto: A multi-sensor state estimator for legged robots in real-world scenarios. *Frontiers in Robotics and AI, 7*. https://doi.org/10.3389/frobt.2020.00068

95. Levine, S., Pastor, P., Krizhevsky, A., & Quillen, D. (2016). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *arXiv preprint arXiv:1603.02199*. https://arxiv.org/abs/1603.02199

96. Achiam, J., Held, D., Tamar, A., & Abbeel, P. (2017). Constrained policy optimization. *arXiv preprint arXiv:1705.10528*. https://arxiv.org/abs/1705.10528

97. Burda, Y., Edwards, H., Storkey, A., & Klimov, O. (2018). Exploration by random network distillation. *arXiv preprint arXiv:1810.12894*. https://arxiv.org/abs/1810.12894

98. Bellemare, M. G., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016). Unifying count-based exploration and intrinsic motivation. *arXiv preprint arXiv:1606.01868*. https://arxiv.org/abs/1606.01868

99. Arena, P., Li Noce, A., & Patanè, L. (2024). Stability and safety learning methods for legged robots. *Robotics, 13*(1), 17. https://doi.org/10.3390/robotics13010017

100. Yang, T.-Y., Zhang, T., Luu, L., Ha, S., Tan, J., & Yu, W. (2022). Safe reinforcement learning for legged locomotion. *arXiv preprint* arXiv:2203.02638. https://arxiv.org/abs/2203.02638

101. Tong, Y., Liu, H., & Zhang, Z. (2024). Advancements in humanoid robots: A comprehensive review and future prospects. *IEEE/CAA Journal of Automatica Sinica, 11*(2), 301–328. https://doi.org/10.1109/JAS.2023.124140

102. Ha, S., Lee, J., van de Panne, M., Xie, Z., Yu, W., & Khadiv, M. (2024). Learning-based legged locomotion: State of the art and future perspectives. *arXiv preprint* arXiv:2406.01152. https://arxiv.org/abs/2406.01152

103. Gu, Z., Li, J., Shen, W., Yu, W., Xie, Z., McCrory, S., Cheng, X., Shamsah, A., Griffin, R., Liu, C. K., Kheddar, A., Peng, X. B., Zhu, Y., Shi, G., Nguyen, Q., Cheng, G., Gao, H., & Zhao, Y. (2025). Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and

learning. *arXiv preprint* arXiv:2501.02116. https://arxiv.org/abs/2501.02116

104. Kang, D., Cheng, J., Zamora, M., Zargarbashi, F., & Coros, S. (2023). RL + model-based control: Using on-demand optimal control to learn versatile legged locomotion. *IEEE Robotics and Automation Letters, 8*(10), 6619–6626. https://doi.org/10.1109/LRA.2023.3307008

105. Margolis, G. B., & Agrawal, P. (2022). Walk these ways: Tuning robot control for generalization with multiplicity of behavior. *arXiv preprint* arXiv:2212.03238. https://arxiv.org/abs/2212.03238

106. Gonzalez, C., Barasuol, V., Frigerio, M., Featherstone, R., Caldwell, D. G., & Semini, C. (2020). Line walking and balancing for legged robots with point feet. *arXiv preprint* arXiv:2007.01087. https://arxiv.org/abs/2007.01087

107. He, J., Chong, B., Xu, Z., Ha, S., & Goldman, D. I. (2024). Learning to enhance multi-legged robot on rugged landscapes. *arXiv preprint* arXiv:2409.09473. https://arxiv.org/abs/2409.09473

108. Denys Kotelovych, Yurii Borzov та Ihor Malets. “Stabilization of a Quintuple Inverted Pendulum System in Isaac Sim”. В: Lecture Notes in Data Engineering, Computational Intelligence, and Decision-Making, Volume 1. За ред. Sergii Babichev та Volodymyr Lytvynenko. Cham: Springer Nature Switzerland, 2024, с. 178—193. ISBN: 978-3-031-70959-3. DOI: 10.1007/978-3-031-70959-3_9.



**Оглядова стаття.**
Надійшла до редакції 20.10.2025.
Прийнято до публікації 26.11.2025.