American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)

E-ISSN: XXXX-XXXX • Website: www.ajamri.com • Email: editor@ajamri.com

Robotic Intelligence in Unstructured Environments: Learning, Perception and Decision-Making Beyond Controlled Settings

Dr. Adrian Lucas Moretti

Associate Professor, University of Bologna, Bologna, Italy

Abstract

Robots designed for controlled factories generally operate within predictable spatial arrangements, stable lighting, standardized objects, and carefully separated human work zones. Unstructured environments present a fundamentally different challenge. Agricultural fields, disaster locations, construction sites, forests, mines, homes, oceans, and planetary surfaces contain irregular terrain, deformable objects, incomplete maps, environmental disturbances, uncertain sensor observations, and dynamic interaction with people or animals. A robot operating in such conditions must do more than repeat a programmed trajectory. It must perceive incomplete evidence, estimate uncertainty, learn from variation, select safe actions, and recover when its assumptions fail.

This study develops a simulation-based framework for comparing five robotic-intelligence architectures: rule-based control, learning-enabled perception, end-to-end learned policy, multimodal uncertainty-aware robotics, and adaptive hybrid robotic intelligence. The architectures were evaluated through perception robustness, environmental generalization, decision quality, recovery capability, safety, and operational feasibility. Their simulated composite reliability scores were 52, 65, 73, 84, and 90, respectively. These values are illustrative scenario outputs rather than empirical robot-test results. The analysis indicates that learning improves performance beyond fixed rules but does not independently guarantee safe generalization. Stronger performance emerges when learning is combined with multimodal perception, explicit uncertainty estimation, model-based safety constraints, online adaptation, and fallback behavior.

The study concludes that field-ready robotic intelligence should be designed as a layered adaptive system rather than a single end-to-end model. Learned perception and control can expand flexibility, while classical estimation, predictive planning, safety filters, and human supervision provide structure and recoverability. The resulting architecture should recognize unfamiliar conditions, reduce operational authority when confidence declines, and collect evidence for later improvement. Progress beyond controlled settings will therefore depend not only on larger datasets and more capable models but also on transparent evaluation, simulation-to-reality validation, resilient hardware, and safety-centered deployment governance.

Keywords: robotic intelligence; unstructured environments; robot learning; multimodal perception; adaptive control; autonomous decision-making; field robotics; uncertainty-aware robotics

American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)

E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

1. Introduction

Industrial robotics achieved its early success by reducing environmental uncertainty. Factory work cells are designed around fixed coordinates, standardized components, repeatable motions, controlled illumination, and physical barriers. A robot can perform the same movement thousands of times because both the task and the surrounding world are intentionally stabilized. This approach has produced high precision and productivity, but it does not transfer directly to environments that cannot be engineered around the machine.

Unstructured environments contain variation that is difficult to enumerate in advance. An agricultural robot may encounter plants of different sizes, wet soil, dust, weeds, moving workers, and partially hidden fruit. A search-and-rescue robot may confront unstable debris, smoke, collapsed structures, poor communication, and incomplete maps. A household robot must handle objects that vary in shape, position, texture, fragility, and ownership. Field robots also experience changing weather, variable traction, sensor contamination, mechanical wear, and unexpected human behavior.

These conditions transform robotics from repeated motion execution into embodied decision-making under uncertainty. A capable robot must determine what is present, where it is located, which parts of the environment are traversable or manipulable, how the situation may change, and which action provides an acceptable balance between progress and safety. Errors in perception propagate into planning and control. A misclassified surface can produce a dangerous route, while an inaccurate object pose can lead to a failed grasp or collision.

Machine learning has expanded robotic perception and control. Deep neural networks can identify objects, segment terrain, estimate depth, and learn policies from demonstration, reinforcement, or self-supervised interaction. Simulation enables extensive training without exposing hardware or people to every failure. Domain randomization varies appearance, friction, mass, delay, and other properties to reduce overfitting to one simulated world. Yet learned systems may remain brittle when operational conditions differ from their training distribution.

The central research problem is therefore not simply how to make a robot learn. It is how to combine learning, perception, planning, control, uncertainty estimation, and safety into an intelligence architecture that remains useful when the world violates its assumptions. This paper compares representative architectures through a transparent simulation and proposes a layered approach for operation beyond controlled settings. No real robot experiment is claimed; all numerical results are explicitly illustrative.

2. Review of Literature

2.1 Learning and Transfer Beyond Simulation

Reinforcement learning enables a robot to improve behavior through interaction by maximizing expected cumulative reward. In a Markov decision process, a policy $\pi(a|s)$ selects action a from state s.

The objective can be written as:

$J(\pi)=E\pi[t=0\sum T\gamma tr(st,at)]$

where r is the reward, γ is the discount factor, and T is the decision horizon. The formulation is powerful because it can optimize sequential behavior without requiring every control rule to be specified manually.

Real-world learning is nevertheless expensive and risky. Exploratory actions can damage equipment, environments, or people. Simulation provides a safer alternative, but policies trained in simulation often perform poorly on physical systems because models cannot perfectly reproduce contact dynamics, sensor noise, latency, wear, or environmental variation. This discrepancy is commonly described as the simulation-to-reality gap.

Tobin and colleagues demonstrated that domain randomization could support transfer by exposing a model to varied simulated appearances. Peng and colleagues extended randomization to physical dynamics, including mass and friction. Muratore and colleagues examined domain randomization for policy optimization, while Mehta and colleagues proposed active domain randomization that concentrates training on informative variations rather than sampling all parameters uniformly. These studies show that diversity can improve generalization, although excessively broad randomization may produce conservative or unstable policies.

2.2 Multimodal Perception and Environmental Understanding

Robotic perception ordinarily combines cameras, lidar, radar, inertial sensors, joint encoders, force sensors, tactile arrays, microphones, or positioning systems. Each modality has different failure characteristics. Cameras provide semantic richness but are affected by lighting, smoke, glare, and occlusion. Lidar provides geometric structure but may be degraded by dust, rain, or reflective surfaces. Tactile sensing becomes informative only after contact but can reveal properties that vision cannot reliably infer.

Multimodal fusion can reduce dependence on any single sensor. The system may integrate measurements at the raw-data, feature, state-estimation, or decision level. An effective fusion architecture should account for confidence rather than treating every signal as equally reliable. If a camera becomes obscured, the robot should reduce its influence and rely more heavily on geometry, proprioception, or touch.

Perception in unstructured environments requires more than object labeling. The robot must estimate traversability, affordances, motion, uncertainty, and relationships among objects. A fallen branch may be an obstacle for one robot, a manipulable object for another, and a support structure for a climbing robot. Perception must therefore be connected to embodiment and task capability.

2.3 Planning, Adaptation, and Safe Autonomy

Classical motion planning provides structured ways to search for collision-free trajectories, while model-predictive control repeatedly optimizes actions over a moving horizon. Learning-based policies can respond rapidly and represent complex behavior, but their internal decision process may be difficult to verify. Hybrid systems combine learned perception or policy components with explicit dynamics, constraints, or safety controllers.

Uncertainty is central to safe action selection. Aleatoric uncertainty arises from inherent environmental variability or noisy observations, whereas epistemic uncertainty reflects limited knowledge or insufficient training coverage. A robot should distinguish between these forms because additional observations may reduce epistemic uncertainty but not necessarily eliminate environmental randomness.

American Journal of Advanced Multidisciplinary Research and Innovation (AJAMRI)

E-ISSN: XXXX-XXXX ● Website: www.ajamri.com ● Email: editor@ajamri.com

Safe operation may require a robot to slow down, request assistance, select a conservative route, or stop when confidence becomes inadequate. A system that always produces an action can be less intelligent than one capable of recognizing the boundaries of its competence. Research on learning where to trust unreliable models reinforces the importance of identifying states where a learned model is likely to fail.

3. Objectives of the Study

3.1 Technical Objectives

The first objective is to compare representative robotic-intelligence architectures through a common reliability framework. The study evaluates whether learning-based perception and control improve performance beyond fixed rules and whether additional gains arise from multimodal sensing, uncertainty estimation, and hybrid safety mechanisms.

The second objective is to examine how perception, learning, planning, and recovery interact. Field reliability depends on the complete perception–decision–action loop. A highly accurate perception model may provide limited value if the planner cannot respond safely to uncertainty or the robot cannot recover from physical disturbance.

3.2 Deployment and Safety Objectives

The study aims to identify design principles for operating in agriculture, disaster response, construction, domestic service, environmental monitoring, and other variable settings. Particular attention is given to unfamiliar terrain, sensor degradation, communication loss, dynamic obstacles, and environmental change.

A further objective is to develop a responsible basis for reporting field-robotics performance. Laboratory success, simulation success, and real-world reliability should be distinguished. Claims must specify environmental conditions, failure definitions, intervention levels, uncertainty boundaries, and safety procedures.

4. Research Methodology

4.1 Research Design

A simulation-based comparative design was adopted because no original robotic trials or field dataset were supplied. Five architectures were constructed as analytical scenarios: rule-based control, learning-enabled perception, end-to-end learned policy, multimodal uncertainty-aware system, and adaptive hybrid robotic intelligence.

The analysis does not reproduce the performance of a particular commercial robot or published experiment. It synthesizes relationships established in robot-learning, perception, domain-transfer, and safe-control literature. The resulting values demonstrate an assessment method and must not be interpreted as measured product specifications.

4.2 Architecture Definition

Rule-based control represented a conventional system with manually specified states, thresholds, paths, and responses. Learning-enabled perception added machine-learning-based recognition or terrain estimation while retaining conventional planning and control. The end-to-end configuration learned a direct mapping from observation to action.

The multimodal uncertainty-aware architecture fused complementary sensors, estimated predictive uncertainty, and adjusted decisions according to confidence. The adaptive hybrid architecture combined learned perception and policy components with state estimation, model-predictive planning, safety constraints, online adaptation, recovery behavior, and human-directed escalation.

Table 1. Simulated robotic-intelligence architectures

Architecture	Principal capability	Main limitation
Rule-based control	Predictable behavior under known conditions	Limited response to unforeseen variation
Learning-enabled perception	Improved recognition and terrain interpretation	Downstream control remains inflexible
End-to-end learned policy	Direct adaptive observation-to-action mapping	Limited interpretability and safety assurance
Multimodal uncertainty-aware system	Sensor fusion and confidence-sensitive decisions	Higher computational and integration burden
Adaptive hybrid intelligence	Learning, planning, safety filtering, recovery, and adaptation	Complex validation and maintenance requirements

Note: Architectures are conceptual and not tied to a named robotic platform.

4.3 Evaluation Criteria

Six criteria were used: perception robustness, environmental generalization, decision quality, recovery capability, safety, and operational feasibility. Each criterion was scored on a five-point scale. The **composite field-robotics reliability index was defined as:**

$FRRIj=100(0.20Pj+0.18Gj+0.18Dj+0.15Rj+0.19Sj+0.10Fj)/5$

where FRRIj is the index for architecture j; P represents perception robustness; G, environmental generalization; D, decision quality; R, recovery; S, safety; and F, operational feasibility.

Perception and safety received relatively high weights because failures in these dimensions can propagate through the complete robotic system. Operational feasibility remained included because an architecture that cannot meet latency, energy, maintenance, or hardware constraints may not be deployable despite strong theoretical performance.

4.4 Scenario Construction and Analytical Procedure

The simulated scenario involved a mobile manipulation robot operating in a partially mapped disaster environment. Conditions included irregular debris, poor illumination, dust, moving responders, intermittent communication, uncertain traction, and previously unseen objects. The robot was required to navigate, inspect, manipulate lightweight obstacles, and return safely.

The configurations were scored through literature-informed directional assumptions. Weighted values were aggregated and compared. No inferential statistics were used because the scores were constructed rather than sampled. Sensitivity was considered by examining whether the ranking would change if operational feasibility received greater weight.

5. Analysis

5.1 Comparative Reliability

Rule-based control achieved a simulated score of 52. It offered predictable behavior when conditions matched predefined rules but performed weakly when terrain, objects, or sensor observations departed from expected patterns. Its limited capacity to recover from unmodeled situations reduced overall reliability.

Learning-enabled perception achieved 65. Improved recognition supported more accurate terrain and object interpretation, but perception gains did not eliminate limitations in planning and control. The end-to-end learned policy achieved 73 because it could represent complex responses without manually encoding every condition. However, limited interpretability, unpredictable out-of-distribution behavior, and weak safety assurance constrained its score.

The multimodal uncertainty-aware system achieved 84. Its stronger performance arose from sensor redundancy and confidence-sensitive action selection. Adaptive hybrid intelligence achieved the highest score of 90 because it combined flexible learning with structured planning, safety filtering, recovery, and online adaptation.

Table 2. Simulated reliability assessment

Architecture	Perception robustness	Generalization	Recovery and safety	Reliability index
Rule-based control	2.5/5	1.8/5	2.8/5	52
Learning-enabled perception	3.8/5	3.0/5	3.1/5	65
End-to-end learned policy	3.9/5	3.8/5	3.2/5	73
Multimodal uncertainty-aware system	4.5/5	4.2/5	4.3/5	84
Adaptive hybrid robotic intelligence	4.7/5	4.6/5	4.7/5	90

Source: Author-developed simulation informed by established robotics literature.
Caution: Values are illustrative and do not represent physical robot-test results.

5.2 Perception and Decision Coupling

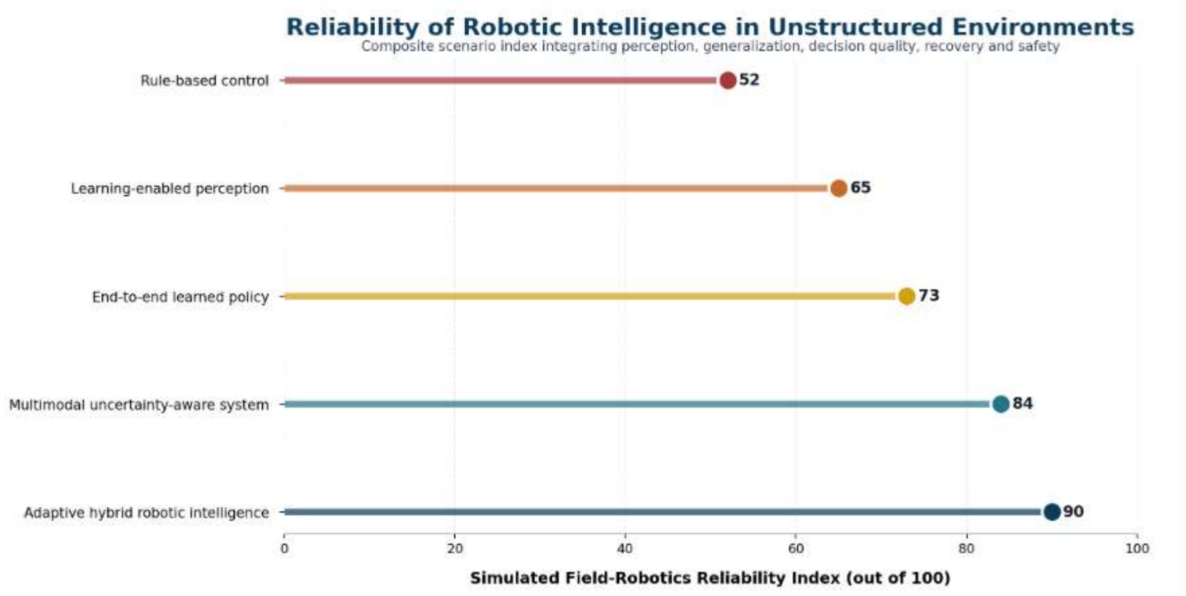
The simulated comparison indicates that perception accuracy cannot be evaluated separately from decision consequences. A terrain-classification error is more serious when it causes the robot to enter an unstable area than when it merely delays progress. Perception metrics should therefore be connected to navigation risk, grasp success, collision probability, and task completion.

systems provide greater resilience when individual sensors fail in different ways. Their value depends on reliable synchronization, calibration, confidence estimation, and failure detection. Fusion can reduce error only when the system recognizes corrupted or misleading signals. Otherwise, an inaccurate sensor may contaminate the combined state estimate.

Adaptive hybrid intelligence achieved the strongest result because it coupled learned representations with explicit action constraints. When uncertainty was low, the learned policy could support rapid progress. When uncertainty increased, the system could slow down, gather additional observations, change sensors, select a conservative path, or request human assistance.

5.3 Graphical Presentation

Graph 1. Reliability of Robotic Intelligence in Unstructured Environments



Interpretation: Learning-enabled components progressively increase the simulated reliability index. The most substantial improvement appears when multimodal perception and explicit uncertainty management are added. Hybrid architecture produces the highest score by combining learned flexibility with structured safety and recovery mechanisms.

Key finding: Adaptive hybrid robotic intelligence scored 38 points above rule-based control and 17 points above the end-to-end learned policy, suggesting that field reliability depends on coordinated sensing, planning, adaptation, and safety rather than learning alone.

Source: Author-developed simulation. Values are illustrative and are not empirical robot-test results.

6. Challenges

6.1 The Open-World Perception Problem

A robot cannot be trained on every object, terrain condition, weather pattern, or human action it will encounter. Models frequently assume a closed set of categories, while real environments contain unfamiliar objects and ambiguous combinations. A system may assign an unknown observation to the nearest familiar class with unjustified confidence.

Open-set recognition and out-of-distribution detection can help identify unfamiliar inputs, but they do not resolve every case. Environmental novelty may be gradual rather than obvious. A surface can appear visually familiar while having unusual friction or structural weakness. Robots must combine semantic perception with physical interaction and uncertainty-sensitive exploration.

6.2 Simulation-to-Reality Transfer

Simulators simplify contact, deformation, lighting, material properties, actuator response, latency, and sensor noise. A policy can exploit simulation artifacts that do not exist in the physical world. High simulation performance may therefore provide weak evidence of real-world safety.

Domain randomization improves exposure to variation, but choosing realistic parameter ranges remains difficult. Excessive randomization may generate physically implausible worlds, while narrow ranges may fail to prepare the system for actual uncertainty. Transfer evaluation should include controlled laboratory tests, progressively difficult field trials, and rare-event scenarios.

6.3 Safe Learning and Exploration

Reinforcement learning often depends on trial and error. In physical environments, unsafe trials can injure people, damage equipment, or disturb ecosystems. Reward functions may also produce unintended behavior when they fail to represent operational goals completely.

Safe exploration requires constraints, shielding, human supervision, or validated simulation. The robot should distinguish between reversible low-risk experiments and actions with serious consequences. Learning should never be permitted to override emergency stops, force limits, protected zones, or other essential safeguards.

6.4 Hardware, Energy, and Communication Constraints

Complex perception and planning models require processing power, memory, energy, and thermal management. Field robots may have limited battery capacity and unreliable communication. Cloud-based processing can increase capability but introduces latency, connectivity, privacy, and cybersecurity concerns.

Robotic intelligence must therefore be co-designed with hardware. Models may need compression, event-driven operation, hierarchical control, or selective sensor activation. A system that performs strongly on a workstation may be unsuitable for an autonomous platform with strict energy and latency limits.

7. Discussion

7.1 Beyond End-to-End Intelligence

End-to-end learning is attractive because it reduces the need for manually engineered intermediate representations. However, an opaque mapping from observation to action creates challenges for diagnosis and assurance. When a robot fails, developers need to determine whether the cause was perception, state estimation, planning, control, or hardware.

A modular system supports diagnosis but may suffer from accumulated interface errors. Hybrid architecture offers a middle path: learned components handle complex perception and adaptation, while explicit estimation, planning, and safety mechanisms preserve structure. The appropriate balance depends on task speed, environmental variation, and consequences of error.

The study's simulated results support this hybrid approach. Learning expanded flexibility, but multimodal uncertainty estimation and recovery produced larger reliability gains. The most capable robot was not the system with the fewest engineered components; it was the system able to shift between learned behavior, explicit planning, conservative control, and human assistance.

7.2 Uncertainty as an Operational Signal

Uncertainty should not remain a diagnostic number displayed after a decision. It should influence how the robot acts. High uncertainty may trigger slower motion, wider obstacle clearance, additional sensing, alternative viewpoints, or human consultation. Low confidence should reduce autonomy rather than merely accompany an unchanged action.

Calibration is important because a confidence value is useful only when it corresponds reasonably to observed correctness. Overconfident robots may enter hazardous states, while excessively conservative robots may fail to complete useful work. Calibration should be evaluated across environmental conditions rather than only on aggregated test data.

Uncertainty also provides a mechanism for continual learning. Difficult cases can be stored for expert labeling, simulation reconstruction, or later training. However, online learning must be controlled to prevent corrupted observations or isolated events from degrading established behavior.

7.3 Deployment and Governance Implications

Field deployment should proceed through staged autonomy. Robots can begin with supervised operation, move to shared control, and receive greater authority only after demonstrating reliable performance across representative conditions. The process should include predefined safety thresholds, intervention criteria, and rollback procedures.

Evaluation must report more than average task success. Relevant measures include collision rate, near misses, human interventions, recovery success, uncertainty calibration, energy consumption, latency, environmental disturbance, and performance under sensor failure. Rare but severe failures require particular attention.

Responsibility must remain clear. Manufacturers, software developers, operators, system integrators, and deploying organizations should understand their respective obligations. Incident logs and version records are essential because learning-enabled behavior may change as models, data, or environmental conditions evolve.

8. Conclusion

Robotic intelligence in unstructured environments requires capabilities that extend far beyond repeated control. Robots must interpret incomplete observations, generalize beyond training conditions, select actions under uncertainty, and recover when predictions fail. Learning contributes important flexibility, but it does not independently establish field reliability.

The simulation-based comparison showed progressive improvement from fixed rules to adaptive hybrid intelligence. Learning-enabled perception improved environmental interpretation, while end-to-end policies expanded behavioral flexibility. Multimodal uncertainty-aware architecture provided a stronger increase by supporting sensor redundancy and confidence-sensitive action. The adaptive hybrid system achieved the highest simulated reliability because it integrated learning with planning, constraints, recovery, and controlled adaptation.

The numerical findings are illustrative and cannot substitute for real field experimentation. Future empirical studies should evaluate representative environmental diversity, physical disturbances, sensor degradation, communication loss, unfamiliar objects, and human interaction. Performance should be reported with repeated trials, uncertainty intervals, failure taxonomies, intervention rates, and clear operational boundaries.

The principal conclusion is that robots will move successfully beyond controlled settings only when intelligence is designed as a layered embodied capability. Perception must inform action, learning must respect safety, uncertainty must change behavior, and failure must lead to recovery rather than uncontrolled continuation. Such systems can support agriculture, infrastructure inspection, disaster response, environmental monitoring, and other socially valuable applications while maintaining defensible limits on autonomous authority.

References

1. Abbeel, P., Coates, A., & Ng, A. Y. (2010). Autonomous helicopter aerobatics through apprenticeship learning. *The International Journal of Robotics Research, 29*(13), 1608–1639. https://doi.org/10.1177/0278364910371999
2. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv*. https://doi.org/10.48550/arXiv.1606.06565
3. Bousmalis, K., Irpan, A., Wohlhart, P., Bai, Y., Kelcey, M., Kalakrishnan, M., Downs, L., Ibarz, J., Pastor, P., Konolige, K., Levine, S., & Vanhoucke, V. (2018). Using simulation and domain adaptation to improve efficiency of deep robotic grasping. In *2018 IEEE International Conference on Robotics and Automation* (pp. 4243–4250). https://doi.org/10.1109/ICRA.2018.8460875
4. Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal on Robotics and Automation, 2*(1), 14–23. https://doi.org/10.1109/JRA.1986.1087032
5. Calandra, R., Owens, A., Jayaraman, D., Lin, J., Yuan, W., Malik, J., Adelson, E. H., & Levine, S. (2018). More than a feeling: Learning to grasp and regrasp using vision and touch. *IEEE Robotics and Automation Letters, 3*(4), 3300–3307. https://doi.org/10.1109/LRA.2018.2852779
6. Deisenroth, M. P., Neumann, G., & Peters, J. (2013). *A survey on policy search for robotics*. Now Publishers.

7. Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the 33rd International Conference on Machine Learning* (pp. 1050–1059).
8. Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proceedings of the 35th International Conference on Machine Learning* (pp. 1861–1870).
9. Kadian, A., Truong, J., Gokaslan, A., Clegg, A., Wijmans, E., Lee, S., Savva, M., Chernova, S., & Batra, D. (2020). Sim2Real predictivity: Does evaluation in simulation predict real-world performance? *IEEE Robotics and Automation Letters, 5*(4), 6670–6677. https://doi.org/10.1109/LRA.2020.3013848
10. Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., Quillen, D., Holly, E., Kalakrishnan, M., Vanhoucke, V., & Levine, S. (2018). QT-Opt: Scalable deep reinforcement learning for vision-based robotic manipulation. *arXiv*. https://doi.org/10.48550/arXiv.1806.10293
11. Kendall, A., & Gal, Y. (2017). What uncertainties do we need in Bayesian deep learning for computer vision? In *Advances in Neural Information Processing Systems* (pp. 5574–5584).
12. Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research, 32*(11), 1238–1274. https://doi.org/10.1177/0278364913495721
13. Levine, S., Pastor, P., Krizhevsky, A., Ibarz, J., & Quillen, D. (2018). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *The International Journal of Robotics Research, 37*(4–5), 421–436. https://doi.org/10.1177/0278364917710318
14. Mehta, B., Diaz, M., Golemo, F., Pal, C. J., & Paull, L. (2020). Active domain randomization. In *Proceedings of the Conference on Robot Learning* (pp. 1162–1176).
15. Muratore, F., Treede, F., Gienger, M., & Peters, J. (2018). Domain randomization for simulation-based policy optimization with transferability assessment. In *Proceedings of the Conference on Robot Learning* (pp. 700–713).
16. Peters, J., & Schaal, S. (2008). Reinforcement learning of motor skills with policy gradients. *Neural Networks, 21*(4), 682–697. https://doi.org/10.1016/j.neunet.2008.02.003
17. Pinto, L., Davidson, J., Sukthankar, R., & Gupta, A. (2017). Robust adversarial reinforcement learning. In *Proceedings of the 34th International Conference on Machine Learning* (pp. 2817–2826).
18. Punjani, A., & Abbeel, P. (2015). Deep learning helicopter dynamics models. In *2015 IEEE International Conference on Robotics and Automation* (pp. 3223–3230).
19. Sadeghi, F., & Levine, S. (2017). CAD2RL: Real single-image flight without a single real image. In *Robotics: Science and Systems*. https://doi.org/10.15607/RSS.2017.XIII.034
20. Sünderhauf, N., Brock, O., Scheirer, W., Hadsell, R., Fox, D., Leitner, J., Upcroft, B., Abbeel, P., Burgard, W., Milford, M., & Corke, P. (2018). The limits and potentials of deep learning for robotics. *The International Journal of Robotics Research, 37*(4–5), 405–420. https://doi.org/10.1177/0278364918770733
21. Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic robotics*. MIT Press.
22. 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. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 23–30). https://doi.org/10.1109/IROS.2017.8202133