

**ИССЛЕДОВАНИЕ АЛГОРИТМОВ ПОВЫШАЮЩИХ КАЧЕСТВО  
ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ В СИМУЛЯЦИИ ДЛЯ ИСПОЛЬЗОВАНИЯ  
НА РЕАЛЬНЫХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМАХ**

**Труфанова А.А** (Университет ИТМО), **Симонов Р.А** (Университет ИТМО),  
**Симонов Н.А** (Университет ИТМО)

**Научный руководитель – к.т.н., доцент Ведяков А.А.** (Университет ИТМО)

В представленной работе исследованы существующие подходы переноса предобученных моделей на реальные робототехнические системы — задача sim-to-real. Были рассмотрены классы подходов, выбрано два из них для тестирования: Domain Randomization and Adaptation и Meta Learning. Апробация производилась на шагающем роботе Unitree A1 в задаче передвижения по сложной поверхности.

**Введение.** На текущий момент разработано и протестировано множество методов машинного обучения для робототехнических систем, использующих нейронные сети для компьютерного зрения [1], имитационное обучение для захвата [2] и походки [3], обучение с подкреплением в задачах автономного управления [4] и многие другие. Большинство из них показывают весьма хорошие результаты обучения в средах моделирования, в то же время полученные модели зачастую не всегда применимы в реальной физической среде. Это происходит из-за невозможности полного переноса физической модели окружения и агента в симуляцию, создания полного аналитического описания, также нередко не учитываются шумы измерений и погрешности датчиков. Для решения проблемы переноса результатов из симуляционной среды в физическую появилась область задач sim-to-real или переноса обученной модели на реальные системы.

**Основная часть.** Большинство методов sim-to-real можно отнести к одному из следующих подходов [5]: Domain Randomization, Transfer Learning, Robust RL, Meta Learning. В них добавляются шумы сигналов датчиков состояний, вводятся дополнительные состояния агента или системы, в архитектуру добавляются новые агенты, обученные на воссоздание помех и динамических возмущений близких к реальным, выполняются разбиения на меньшие по размерности задачи. Многие методы также содержат в себе совокупность нескольких подходов и дополнительное введение сверточных сетей для оптимизации параметров. В исследовании более подробно рассматриваются подходы Domain Randomization and Meta Learning: Domain Randomization and Adaptation [6] кинематических и динамических параметров с использованием Multi-Policy Bayesian Optimization (MPBO). А также алгоритма Meta Strategy Optimization (MSO) [7] подхода мета-обучения.

Для тестирования этих алгоритмов была выбрана задача передвижения по сложной поверхности четырехногого шагающего робота Unitree A1. Под сложной понимается неоднородная поверхность, имеющая уклоны, неровности, перепады высот и тд. Для работы с моделью был выбран физический симулятор PyBullet за счет большого количества поддерживаемых визуализаторов, а также открытого исходного кода. Системе доступны измерения одометрии робота (положения лап относительно системы координат тела, измерения датчиков инерциального блока, расположенного в теле робота), линейные и угловые скорости лап, измерения датчика касания поверхности. Выходом обученной модели является вектор углов для каждого двигателя системы, возвращаемый с заданной частотой. Архитектура модели строится по аналогии с предложенной командой из Гонконгского университета и Baidu Inc. в работе [8].

**Выводы.** В работе разобраны и протестированы методы MPBO и MSO на шагающем роботе Unitree A1, сперва с использованием среды PyBullet, а после на реальной системе. Для обучения

была взята задача передвижения по сложной поверхности. По результатам апробации были составлены рекомендации по использованию исследованных алгоритмов.

**Список использованных источников:**

1. Davies, E.R. and Turk, M., Advanced Methods and Deep Learning in Computer Vision, 2021, AElsevier Science, Computer Vision and Pattern Recognition
2. H. Ben Amor, O. Kroemer, U. Hillenbrand, G. Neumann and J. Peters, Generalization of human grasping for multi-fingered robot hands, 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura-Algarve, Portugal, 2012, pp. 2043-2050
3. X. B. Peng, E. Coumans, T. Zhang, Tsang-Wei E. Lee, J. Tan, S. Levine, Learning Agile Robotic Locomotion Skills by Imitating Animals, 2020, CORL
4. B. R. Kiran et al., Deep Reinforcement Learning for Autonomous Driving: A Survey, in IEEE Transactions on Intelligent Transportation Systems, 2021, vol. 23, no. 6, pp. 4909-4926
5. W. Zhao, J. P. Queralta and T. Westerlund, Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a Survey, 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, ACT, Australia, 2020, pp. 737-744
6. Exarchos, Ioannis and Jiang, Yifeng and Yu, Wenhao and Liu, C. Karen, Policy Transfer via Kinematic Domain Randomization and Adaptation, 2021 IEEE International Conference on Robotics and Automation (ICRA), 2021
7. W. Yu et al., Learning Fast Adaptation with Meta Strategy Optimization, IEEE Robotics and Automation Letters, 2020, pp. 2950–2957
8. H. Shi et al., Reinforcement Learning With Evolutionary Trajectory Generator: A General Approach for Quadrupedal Locomotion, IEEE Robotics and Automation Letters, 2021, vol. 7, no. 2, pp. 3085-3092