

ИССЛЕДОВАНИЕ ПРИМЕНЕНИЯ ДИФФУЗИОННЫХ МОДЕЛЕЙ В РОБОТОТЕХНИЧЕСКИХ ЗАДАЧАХ

Лалаянц К.А. (Университет ИТМО), Дюжев В.Д. (Университет ИТМО)

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

Введение. С каждым годом, методы машинного обучения помогают решить все большее количество задач с непреклонно растущим качеством. Подходы и архитектуры, применяемые в этой области, становятся все более универсальными. Робототехника – одна из наиболее актуальных сфер для их применения, так как классические методы не всегда позволяют автономным системам выполнять свои задачи в динамических окружениях.

Диффузионные модели, так же известные как DDPM [3], IDDPM [4], демонстрируют свои возможности наиболее ярко в обработке изображений, где стали стандартом качества [5, 6]. В последствии, авторы исследований [1, 2] смогли успешно перенести гибкость данного подхода на задачи робототехники. Целью этой работы является исследование указанных методов и выработка методических указаний, для последующего применения в новых задачах.

Основная часть. Методы обучения нейронных сетей, представленные в работах [1, 2], позволяют создать агента на основе нейронных сетей, способного выполнять задачи в динамических окружениях. Модель должна предсказывать набор действий на некоторое количество шагов вперед, основываясь на информации о прошлых состояниях. Обучение проходит на наборе данных, состоящем из пар состояние-действие.

DDPM

В основе предложенных методов лежит использование вероятностных моделей шумоподавления (DDPM), представленных в статье [3, 4], исторически использовавшихся для генерации изображений. Данные для обучения получаются путем поэтапного добавления шумов на действия из обучающей выборки. После чего модель должна предсказать шум, основываясь на входных состояниях. Обученная модель способна из вектора случайно сгенерированных шумов и полученного на вход состояния среды получить вектор действия для решения поставленной задачи.

Diffusion-QL

В исследовании [2] ставится цель связать классический для робототехники метод обучения с подкреплением с вероятностными моделями шумоподавления. Научная новизна данного исследования заключается в предложении использовать диффузионную модель для представления политики в офлайн обучении с подкреплением (Offline RL). Авторы представляют Diffusion Q-learning (Diffusion-QL), использующий условную диффузионную модель для клонирования поведения и регуляризации политики. Результаты показывают эффективность Diffusion-QL, превосходящую предыдущие методы.

Diffusion policy

Авторы работы [1] обнаружили, что диффузионный подход дает значительные преимущества при использовании в политике робота, такие как: хорошая обработка мультимодальных распределений действий; совместимость с высокоразмерными пространствами действий; впечатляющая стабильность при обучении. Все это позволило им превзойти показатели SOTA методов обучения в среднем на 46.9%.

В работе также отмечаются недостатки, например, ограничения, связанные с применением подхода имитации поведения. Одна из основных целей этого исследования – найти способ их обойти с помощью объединения с Diffusion-QL, сохранив плюсы двух методов.

Симуляционные среды

Для апробации полученных результатов используются симуляционные среды на основе библиотек gym, MuJoCo и PyBullet. Обучения моделей проводилось с различными наборами гиперпараметров для подбора наилучших конфигураций по скорости и качеству обучения.

Выводы. В работе рассматривались актуальные методы, использующие DDPM архитектуру, являющуюся довольно перспективной на сегодня. Ставилась цель получить универсальный метод решения робототехнических задач, объединив методы Diffusion policy и Diffusion-QL, и, после проведения экспериментов в симуляционных средах, разработаны методические указания по использованию и настройке исследованных методов для применения в робототехнических системах.

Список использованных источников:

1. Chi C. et al. Diffusion policy: Visuomotor policy learning via action diffusion //arXiv preprint arXiv:2303.04137. – 2023.
2. Wang Z., Hunt J. J., Zhou M. Diffusion policies as an expressive policy class for offline reinforcement learning //arXiv preprint arXiv:2208.06193. – 2022.
3. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models //Advances in neural information processing systems. – 2020. – Т. 33. – С. 6840-6851.
4. Nichol A. Q., Dhariwal P. Improved denoising diffusion probabilistic models //International Conference on Machine Learning. – PMLR, 2021. – С. 8162-8171.
5. Rombach R. et al. High-resolution image synthesis with latent diffusion models //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2022. – С. 10684-10695.
6. Ramesh A. et al. Zero-shot text-to-image generation //International Conference on Machine Learning. – PMLR, 2021. – С. 8821-8831.