

КРИВЕНЧУК ЮРІЙ

Національний університет "Львівська політехніка"

<https://orcid.org/0000-0002-2504-5833>e-mail: yurii.p.kryvenchuk@lpnu.ua**ШАВАРСЬКИЙ МАКСИМ**

Національний університет "Львівська політехніка"

<https://orcid.org/0000-0002-1379-3244>e-mail: maksym.a.shavarskyi@lpnu.ua

ПРОЦЕС ЗБОРУ ТА ПІДГОТОВКИ ДАНИХ З ДЕМОНСТРАЦІЙ, ЗРОБЛЕНИХ ЛЮДИНОЮ, ДЛЯ ВИКОНАННЯ ПРОСТИХ МАНІПУЛЯЦІЙНИХ ЗАВДАНЬ ЗА ДОПОМОГОЮ РОБОРУКИ XARM7 ДЛЯ НАВЧАННЯ З ДОПОМОГОЮ BEHAVIOR CLONING АЛГОРИТМУ

Основна мета цього дослідження полягає в тому, щоб заглибитися в тонкощі визначення потрібної структури даних, необхідної для інструктування робота щодо виконання елементарних завдань маніпулювання. Навчання буде відбуватись завдяки Imitation learning, використовуючи алгоритм behavior cloning. Imitation learning надзвичайно корисний для навчання будь-яких роботизованих систем, особливо з багатосуглобними роборуками. Даний підхід дозволяє уникнути написання складних та громіздких евристичних алгоритмів. Евристичні алгоритми часто містять багато гіперпараметрів, які потрібно підбирати емпіричним шляхом. Imitation learning алгоритми дозволяють розробити End-to-End policy із зібраних даних, що спрощує систему та підвищує її надійність. Збір даних відбувався з роборуки xArm7 та Realsense D435 RGB+Depth камери.

Ключові слова: Imitation learning, behavior cloning, роборука, xArm7, End-to-End policy, Realsense D435, RGB, Depth, збір та підготовка даних

KRYVENCHUK YURI, SHAVARSKYI MAKSYM

Lviv Polytechnic National University

COLLECT AND PREPARE DATASET FROM HUMAN-DEMONSTRATIONS OF SIMPLE MANIPULATION TASKS USING XARM7 ROBOTIC ARM FOR TRAINING BEHAVIOUR CLONING ALGORITHM

The goal of the study is to understand what data structure is necessary to teach a robot to perform simple manipulation tasks. Training will take place thanks to Imitation learning, using the behavior cloning algorithm. Imitation learning is extremely useful for training any robotic system, especially with multi-jointed robots. This approach allows you to avoid writing complex and cumbersome heuristic algorithms. Heuristic algorithms often contain many hyperparameters that must be chosen empirically. Imitation learning algorithms allow you to develop an End-to-End policy from the collected data, which simplifies the system and increases its reliability. In the realm of training robotic systems through Imitation Learning, the criticality of data collection cannot be overstated. The efficacy of the Behavior Cloning algorithm hinges on the quality and speed at which data is gathered. Properly curated and swiftly acquired datasets serve as the bedrock for training algorithms that emulate human behavior, allowing the robot to seamlessly replicate intricate manipulations. The efficiency of data gathering is paramount in addressing the challenges posed by the dynamic nature of real-world environments. Swift adaptation to varying conditions and unforeseen scenarios necessitates a data collection process that is not only rapid but also comprehensive. A well-curated dataset not only expedites the training process but also enhances the model's adaptability and robustness in real-world applications. In conclusion, the success of Imitation Learning algorithms, particularly those underpinned by the Behavior Cloning approach, is contingent on the meticulous and expeditious gathering of data. This underscores the importance of streamlined data collection processes, ensuring that the resultant models are not only accurate but also capable of navigating the complexities of diverse and dynamic environments. Fastidious attention to data gathering is imperative to ensure a diverse and representative dataset. The xArm7 robot, in collaboration with the Realsense D435 camera, serves as an ideal synergy for capturing a myriad of movements and scenarios. The precision of data collection directly influences the algorithm's ability to generalize, making it adept at handling a spectrum of manipulation tasks.

Key words: Imitation learning, behavior cloning, robotic arm, xArm7, End-to-End policy, Realsense D435, RGB, Depth, collect and prepare dataset.

Постановка проблеми

Написання будь-якого евристичного алгоритму для роботизованої системи є доволі складним і тривалим процесом. Такий алгоритм потребує багато даних різних типів, які в будь-якому випадку потрібно обробляти через велике зашумлення середовища. Зазвичай, такі алгоритми є дуже чутливими до будь-яких змін навколишнього середовища, що постійно призводить до збільшення програмного коду та ускладнення логіки.

На заміну цьому підходу прийшли нейронні мережі. За допомогою нейронних мереж можливо написати End-to-End policy алгоритм для виконання завдання помірної складності. Наприклад, відкрити двері, підняти та викинути сміття, помити брудний посуд, тощо. Imitation learning ідеально підходить для виконання таких типів завдань. Даний вид навчання з підкріпленням, використовуючи записи рухів людини, може ефективно навчитись виконувати завдання. Звідси і назва алгоритму behavior cloning.

Дана робота пропонує підходи та способи збирання та підготовки даних для будь-яких роботизованих систем на прикладі роборуки xArm7 та Realsense RGB+Depth камери, а також показує оптимальну структуру та об'єм даних для досягнення мінімальних результатів.

Аналіз останніх джерел

У статті [1] щоб вивчити вплив джерела набору даних, вони збрали дані з трьох різних джерел –

Machine Generated, Proficient-Human і Multi-Human.

Machine Generated (MG). Для того, щоб зібрати дані, було розроблено алгоритм RL, взявши контрольні точки агента, які регулярно зберігаються під час навчання та зібрано 300 траєкторій з кожного контрольного пункту. Отже, ці набори даних складаються з сумішей експертних і субоптимальних даних і нагадують набори даних із звичайного офлайн RL.

Proficient-Human (PH) і Multi-Human (MH). Набори даних збираються людьми за допомогою RoboTurk, платформи віддаленої телеоперації. Набори даних PH складаються з 200 демонстрацій:

- зібрані одним досвідченим телеоператором, а набори даних MH складаються з 300 демонстрацій;
- зібрані 6 телеоператорами різної кваліфікації, кожен з яких надав 50 демонстрацій;
- шість телеоператорів були поділені на три групи (“better”, “okay”, “worse”) по 2 відповідно до навичок.

Ці підмножини даних у Multi-Human дані дозволили нам дослідити здатність алгоритмів працювати з людськими даними різної якості. Запропонована структура даних:

- дані включають end-effector (позиція останнього суглобу роборуки),
- gripper fingers (пристрій, що кріпиться до останнього суглобу роборуки для взаємодії із навколишнім середовищем),
- joints (кут повороту кожного суглобу),
- ground-truth object poses (реальні позиції кожного об'єкту) та зображення(камера на останньому суглобі та зовнішня камера).

Є два простори спостереження – «low-dim» і «image». Обидва включають end-effector і gripper fingers, але вони відрізняються лише тим, чи використовується ground-truth object poses (“low-dim”) або чи ця інформація замінюється доступними спостереженнями камери (“image”).

У статті [2] дані збирались виключно з демонстрацій. Система складається з 2 основних роборук, 2 допоміжних, та 3 камер. Допоміжні роборуки керуються за допомогою людини, інші дві повторюють рухи. На основних роборуках встановлені камери, а також на корпусі платформи, щоби спостерігати одночасно за їхніми рухами. Розмір зображень складають (480x640x3). Окрім 3 RGB зображень також зберігаються позиції та прискорення кожного суглоба допоміжних рук. Важливо використовувати позиції суглобів допоміжних роборук, а не основних, оскільки величина прикладеної сили визначається різницею між ними.

У статті [3] система дистанційного керування використовує гарнітуру Oculus VR, яка підключена до бортового комп'ютера робота через USB-кабель і відстежує два портативні контролери. Телеоператор стоїть за роботом і використовує контролери, щоб керувати роботом у зоні прямої видимості від третьої особи. Робот реагує на рух оператора в циклі керування частотою 10 Гц не в реальному часі. Відносно швидке керування замкнутим циклом дозволяє оператору з легкістю демонструвати широкий спектр завдань і швидко втрутитися, якщо робот збирається перейти в небезпечний стан під час автономного виконання.

Наступним кроком необхідно розмістити робота перед столом із 6-15 предметами домашнього вжитку в рандомізованих позах. Збираємо демонстрації людей для 100 заздалегідь визначених завдань, які охоплюють 9 базових навичок, таких як штовхання та підбирання. Потім модель оцінюється на 29 нових завданнях з використанням нового мовного опису або відео цього завдання. Щоб метод добре виконував ці відкладені завдання, він повинен правильно інтерпретувати нову команду завдання та виводити дії, які відповідають цьому завданню.

Структура даних для навчання: RGB камера на останньому суглобі роборуки, кутова позиція кожного суглобу роборуки, кутове прискорення кожного суглобу, а також зусилля у певний момент часу.

Метою роботи є розробка підходу та способу збирання та підготовки даних для будь-яких роботизованих систем на прикладі роборуки xArm7 та Realsense RGB+Depth камери, а також опис оптимальної структури та об'єму даних для досягнення мінімальних результатів.

Структура даних

Проаналізувавши багато статей, було отримано знання про дані, які потрібно зібрати. Отже дані мають мати наступний вигляд: датасет $D = \{Observation(t), Controls(t)\}$. Observations(t) - це те, що спостерігає робот. Observations(t) складається з:

1. RGB зображення (160x120x3)
2. Depth image (160x120)
3. Позиція останнього суглоба робота для 5 останніх кроків. Позиція суглоба складається з X, Y, Z - position, X, Y, Z, W - orientation. Включення короткої історії кінцевих точок дозволяє роботу виводити швидкість і прискорення зі зміни кутів суглобів.
4. Час

Controls(t) - вивід нейронної мережі складається з кутової швидкості $\omega(t)$ та лінійної швидкості $v(t)$.

Збір даних

Проаналізувавши описані вище статті та мої технічні можливості було вирішено використовувати дані, згенеровані евристичним алгоритмом, а також дані з демонстрацій людини за допомогою допоміжної роборуки. Завдання, яке треба буде виконати за допомогою роборуки - підняти кришку туалету.

Першим кроком потрібно налаштувати комунікацію між камерою та основною програмою для отримання зображення. Для комунікації між процесами було використано бібліотеку ZeroMQ. Процеси спілкуються по протоколу Inter Process Communication (IPC). Сериалізація даних відбувається за допомогою

Google Protocol Buffers (Protobuf). Для роботи з Realsense D435 камерою було використано Intel RealSense SDK 2.0. Для роботи з xArm7 було використано xArm-C++-SDK. Після того, як було створено окремі повідомлення, необхідно об'єднати їх у єдине ціле для зручного зберігання та подальшого використання.

По завершенню написання комунікації між усіма вузлами системи, необхідно написати програмний код, який пов'яже дві xArm7 роботуки, щоб людина могла віддалено керувати основною роботукою за допомогою допоміжної. Для цього потрібно написати клієнт-серверну архітектуру комунікації, де сервер - основна рука, а клієнт - допоміжна.

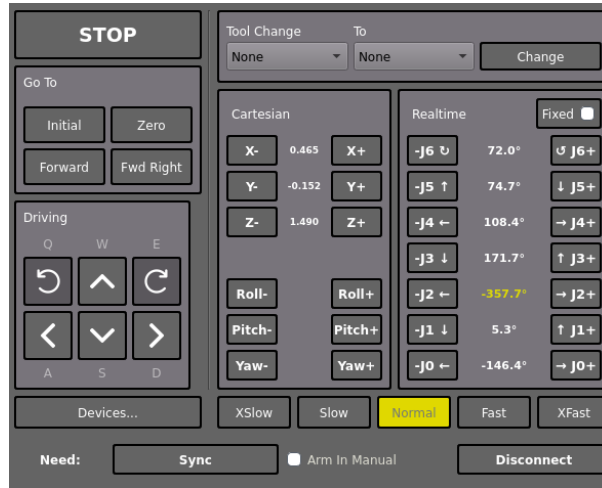


Рис. 1. UI для синхронізації двох xArm7



Рис. 2. Створення демонстрації відкривання кришки туалету основною роботукою за допомогою допоміжної роботуки

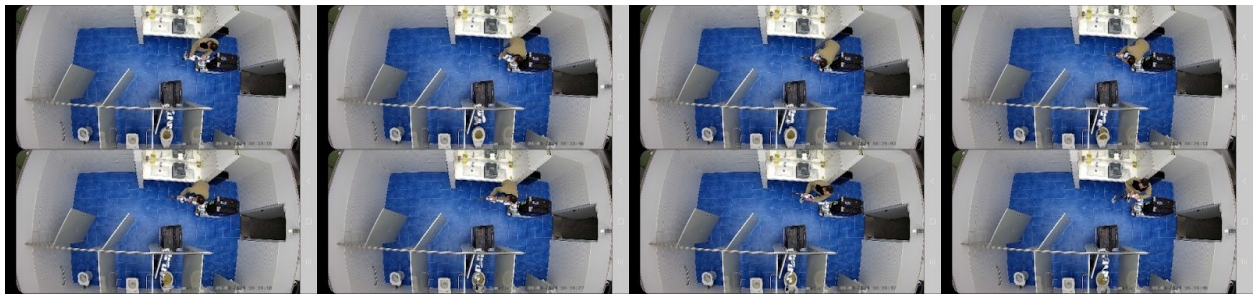


Рис. 3. Створення демонстрації відкривання кришки туалету основною роборукою за допомогою допоміжної роборуки (вид зверху)

Підготовка даних

Для кращого узагальнення моделі, обов'язково необхідно зробити аугментацію даних. Для аугментації даних було використано наступні операції: рандомне віддзеркалення зображення, збільшення/зменшення яскравості, контрасту, hue компоненти (якщо перевести зображення з RGB в HSV формат), додавання шуму та блюр зображення.



Рис. 4. Аугментовані дані для кращої генералізації моделі

Висновок

У роботі зроблено аналіз схожих систем збору та підготовки даних для навчання моделі Imitation learning, а саме Behavior cloning. Запропоновано свій набір даних в залежності від доступних інструментів. Реалізовано комунікацію між Realsense D435 камерою та основною програмою, а також між xArm7 та основною програмою для збору даних. Реалізовано зручний та простий інтерфейс для віддаленого керування основною роборукою за допомогою допоміжної руки. Під час виконання даної роботи було зібрано приблизно 200 демонстрацій за допомогою евристичного алгоритму, середньою тривалістю 15 секунд кожна, та приблизно 100 демонстрацій, зроблений людиною, середньою тривалістю теж 15 секунд. Отже зібрано близько 300 демонстрацій загальною тривалістю 4500 секунд (75 хв). На основі проаналізованих статей, цього повинно вистачити для навчання простої моделі для того щоб підняти кришку туалету.

References

1. Ajay Mandlekar, Danfei Xu, Josiah Wong, Soroush Nasiriany, Chen Wang, Rohun Kulkarni, Li Fei-Fei, Silvio Savarese, Yuke Zhu, Roberto Martín-Martín, "What Matters in Learning from Offline Human Demonstrations for Robot Manipulation", 6 Aug 2021, <https://doi.org/10.48550/arXiv.2108.03298>
2. Tony Z. Zhao, Vikash Kumar, Sergey Levine, Chelsea Finn, "Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware", 4 Jan 2024, <https://doi.org/10.48550/arXiv.2304.13705>
3. Eric Jang, Alex Irpan, Mohi Khansari, Daniel Kappler, Frederik Ebert, Corey Lynch, Sergey Levine, Chelsea Finn, "BC-Z: Zero-Shot Task Generalization with Robotic Imitation Learning", 4 Feb 2022, <https://doi.org/10.48550/arXiv.2202.02005>
4. Christeen T. Jose, "Auxiliary Learning as a step towards Artificial General Intelligence", 30 Nov 2022, <https://doi.org/10.48550/arXiv.2212.00061>
5. Ajay Mandlekar, Danfei Xu, Roberto Martín-Martín, Yuke Zhu, Li Fei-Fei, Silvio Savarese, "Human-in-the-Loop Imitation Learning using Remote Teleoperation", 12 Dec 2020, <https://doi.org/10.48550/arXiv.2012.06733>

-
7. Albert Tung, Josiah Wong, Ajay Mandlekar, Roberto Martín-Martín, Yuke Zhu, Li Fei-Fei, Silvio Savarese, “Learning Multi-Arm Manipulation Through Collaborative Teleoperation”, 12 Dec 2020, <https://doi.org/10.48550/arXiv.2012.06738>
 8. Tianhao Zhang, Zoe McCarthy, Owen Jow, Dennis Lee, Xi Chen, Ken Goldberg, Pieter Abbeel, “Deep Imitation Learning for Complex Manipulation Tasks from Virtual Reality Teleoperation”, 6 Mar 2018, <https://doi.org/10.48550/arXiv.1710.04615>
 9. Liyiming Ke, Jingqiang Wang, Tapomayukh Bhattacharjee, Byron Boots, Siddhartha Srinivasa, “Grasping with Chopsticks: Combating Covariate Shift in Model-free Imitation Learning for Fine Manipulation”, 13 Nov 2020, <https://doi.org/10.48550/arXiv.2011.06719>
 10. RB Ashith Shyam, Zhou Hao, Umberto Montanaro, Gerhard Neumann, “Imitation Learning for Autonomous Trajectory Learning of Robot Arms in Space”, 10 Aug 2020, <https://doi.org/10.48550/arXiv.2008.04007>