Novelty Assessment Report

Paper: Koopman-Assisted Trajectory Synthesis: A Data Augmentation Framework for Offline Imitation Learning

PDF URL: https://openreview.net/pdf?id=UAZCKdd4R7

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-01

Abstract

Data augmentation plays a pivotal role in offline imitation learning (IL) by alleviating covariate shift, yet existing methods remain constrained. Single-step techniques frequently violate underlying system dynamics, whereas trajectory-level approaches are plagued by compounding errors or scalability limitations. Even recent Koopman-based methods typically function at the single-step level, encountering computational bottlenecks due to action-equivariance requirements and vulnerability to approximation errors. To overcome these challenges, we introduce Koopman-Assisted Trajectory Synthesis (KATS), a novel framework for generating complete, multi-step trajectories. By operating at the trajectory level, KATS effectively mitigates compounding errors. It leverages a state-equivariant assumption to ensure computational efficiency and scalability, while incorporating a refined generator matrix to bolster robustness against Koopman approximation errors. This approach enables a more direct and efficacious mechanism for distribution matching in offline IL. Extensive experiments demonstrate that KATS substantially enhances policy performance and achieves state-of-the-art (SOTA) results, especially in demanding scenarios with narrow expert data distributions.

Core Task Landscape

This paper addresses: **Data Augmentation for Offline Imitation Learning**

A total of **50 papers** were analyzed and organized into a taxonomy with **24 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Trajectory-Level Synthesis and Adaptation**
- **Corrective and Interventional Augmentation**
- **Visual and Perceptual Augmentation**
- **Invariance and Equivariance-Based Augmentation**
- **Model-Based Data Generation and Synthesis**
- **Cross-Domain and Transfer-Based Augmentation**
- **Policy-Guided and Expert-Informed Augmentation**
- **Application-Specific Augmentation Methods**
- **Evaluation and Theoretical Foundations**

Complete Taxonomy Tree

- Data Augmentation for Offline Imitation Learning Survey Taxonomy
- Trajectory-Level Synthesis and Adaptation
  - Dynamics-Based Trajectory Generation ★ (5 papers)
  - [0] Koopman-Assisted Trajectory Synthesis: A Data Augmentation Framework for Offline Imitation Learning (Anon et al., 2026) View paper
  - [1] Mimicgen: A data generation system for scalable robot learning using human demonstrations (Mandlekar, 2023) View paper
  - [5] DemoGen: Synthetic Demonstration Generation for Data-Efficient Visuomotor Policy Learning (Xue, 2025) View paper
  - [36] Offline Imitation Learning with Model-based Reverse Augmentation (Jie-Jing Shao, 2024) View paper
  - [40] Offline Trajectory Optimization for Offline Reinforcement Learning (ZIQI ZHAO, 2024) View paper
  - Diffusion-Based Trajectory Synthesis (4 papers)
  - [23] Learning from Random Demonstrations: Offline Reinforcement Learning with Importance-Sampled Diffusion Models (Fang, 2024) View paper
  - [25] DiffStitch: Boosting Offline Reinforcement Learning with Diffusion-based Trajectory Stitching (Li Guanghe, 2024) View paper
  - [30] Diffusion model is an effective planner and data synthesizer for multi-task reinforcement learning (He, 2023) View paper
  - [45] BiTrajDiff: Bidirectional Trajectory Generation with Diffusion Models for Offline Reinforcement Learning (Qing, 2025) View paper
  - Demonstration Adaptation and Replay (2 papers)
  - [4] Variable-speed teachingâplayback as real-world data augmentation for imitation learning (Nozomu Masuya, 2024) View paper
  - [24] Self-evolved Imitation Learning in Simulated World (Ye, 2025) View paper
- Corrective and Interventional Augmentation
  - Continuity-Based Corrective Labeling (2 papers)
  - [2] CCIL: Continuity-based Data Augmentation for Corrective Imitation Learning (Ke, 2023) View paper
  - [3] Guided Data Augmentation for Offline Reinforcement Learning and Imitation Learning (Qu Yuxiao, 2023) View paper
  - Interactive Intervention Generation (2 papers)
  - [7] IntervenGen: Interventional Data Generation for Robust and Data-Efficient Robot Imitation Learning (Ryan Hoque, 2024) View paper

    - [20] Planning in Autonomous Driving Using Imitation Learning With Research on Data Aggregation (Mingjing Liang, 2024) View paper
- Visual and Perceptual Augmentation
    - Novel View Synthesis with NeRF (2 papers)
    - [18] NeRF in the Palm of Your Hand: Corrective Augmentation for Robotics via Novel-View Synthesis (Allan Zhou, 2023) View paper
    - [26] Tube-NeRF: Efficient Imitation Learning of Visuomotor Policies From MPC via Tube-Guided Data Augmentation and NeRFs (Andrea Tagliabue, 2024) View paper
    - State-Conditioned Image Synthesis (2 papers)
    - [35] Imitation Learning through Image Augmentation Using Enhanced Swin Transformer Model in Remote Sensing (Yoojin Park, 2023) View paper
    - [46] S2p: State-conditioned image synthesis for data augmentation in offline reinforcement learning (Cho, 2022) View paper
    - Bimanual and Multi-View Augmentation (2 papers)
    - [15] D-CODA: Diffusion for Coordinated Dual-Arm Data Augmentation (Liu, 2025) View paper
    - [50] ROPA: Synthetic Robot Pose Generation for RGB-D Bimanual Data Augmentation (Chen, 2025) View paper
- Invariance and Equivariance-Based Augmentation
    - SE(3) Equivariant Augmentation (3 papers)
    - [9] RoCoDA: Counterfactual Data Augmentation for Data-Efficient Robot Learning from Demonstrations (Ezra Ameperosa, 2024) View paper
    - [42] Equivariant Data Augmentation for Generalization in Offline Reinforcement Learning (Cristina Pinneri, 2023) View paper
    - [49] SEIL: Simulation-augmented Equivariant Imitation Learning (Mingxi Jia, 2023) View paper
    - Counterfactual and Causal Augmentation (1 papers)
    - [8] Good data is all imitation learning needs (Amir Samadi, 2024) View paper
- Model-Based Data Generation and Synthesis
    - World Model-Based Synthesis (3 papers)
    - [31] Unsupervised Data Generation for Offline Reinforcement Learning: A Perspective from Model (He, 2025) View paper
    - [34] Efficient Imitation Learning with Conservative World Models (Rafailov, 2024) View paper
    - [37] Improving offline reinforcement learning with inaccurate simulators (Yiwen Hou, 2024) View paper
    - Dataset Distillation and Compression (2 papers)
    - [13] Dataset Distillation for Offline Reinforcement Learning (Jonathan Light, 2024) View paper
    - [27] Offline behavior distillation (Shiye Lei, 2024) View paper
    - Synthetic Experience Replay (1 papers)
    - [14] Synthetic experience replay (Lu Cong, 2023) View paper
- Cross-Domain and Transfer-Based Augmentation
    - Simulator-to-Real Transfer (4 papers)
    - [6] Beyond ood state actions: Supported cross-domain offline reinforcement learning (Jinxin Liu, 2024) View paper
    - [19] Synthetic data generation using imitation training (Aman Kishore, 2021) View paper
    - [43] Synthetic Data is Sufficient for Zero-Shot Visual Generalization from Offline Data (Bogunovic, 2025) View paper
    - [48] Synthetic Dataset Generation and Learning From Demonstration Applied to Industrial Manipulation (Alireza Barekatain, 2024) View paper
    - Multi-Task and Meta-Learning Augmentation (1 papers)
    - [11] Disentangling policy from offline task representation learning via adversarial data augmentation (Jia, 2024) View paper
- Policy-Guided and Expert-Informed Augmentation
    - Parametric Expert-Based Augmentation (1 papers)
    - [12] Data augmentation for efficient learning from parametric experts (Galashov, 2022) View paper
    - Expert Identification and Filtering (2 papers)
    - [33] Identifying expert behavior in offline training datasets improves behavioral cloning of robotic manipulation policies (Qiang Wang, 2023) View paper
    - [41] Offline inverse reinforcement learning (Jarboui, 2021) View paper
    - High-Quality Synthetic Data Generation (1 papers)
    - [44] HIPODE: Enhancing Offline Reinforcement Learning with High-Quality Synthetic Data from a Policy-Decoupled Approach (Lian, 2023) View paper
- Application-Specific Augmentation Methods
    - Autonomous Driving Augmentation (1 papers)
    - [17] An imitation learning method with data augmentation and post processing for planning in autonomous driving (W Xi, 2023) View paper
    - Game AI and Generalization (1 papers)
    - [16] Improving Generalization in Game Agents with Data Augmentation in Imitation Learning (Derek Yadgaroff, 2023) View paper
    - Specialized Domain Applications (3 papers)
    - [10] DABI: Evaluation of Data Augmentation Methods Using Downsampling in Bilateral Control-Based Imitation Learning with Images (Masato Kobayashi, 2024) View paper
    - [21] Towards imitation learning to branch for mip: A hybrid reinforcement learning based sample augmentation approach (C Zhang, 2024) View paper
    - [29] RL-Fill: Timing-Aware Fill Insertion Using Reinforcement Learning (Jinoh Cho, 2024) View paper
- Evaluation and Theoretical Foundations
    - Augmentation Effectiveness Analysis (2 papers)
    - [32] Value Function Evaluation with Data Augmentation for Offline Reinforcement Learning (Xianwei Zhou, 2024) View paper
    - [47] Leveraging Haptic Feedback to Improve Data Quality and Quantity for Deep Imitation Learning Models. (Catie Cuan, 2025) View paper
    - Horizon and Behavioral Analysis (2 papers)
    - [22] Is behavior cloning all you need? understanding horizon in imitation learning (Adam Block, 2024) View paper
    - [28] Offline imitation learning through graph search and retrieval (Zhao-Heng Yin, 2024) View paper
    - Reverse Model and Optimization Studies (2 papers)

- [38] Efficient Deep Learning of Robust Policies From MPC Using Imitation and Tube-Guided Data Augmentation (Andrea Tagliabue, 2023) View paper
- [39] ROMA: Reverse Model-Based Data Augmentation for Offline Reinforcement Learning (Xiaochen Wei, 2023) View paper

Narrative

Core task: data augmentation for offline imitation learning. The field addresses the challenge of learning policies from fixed demonstration datasets by generating additional training data. The taxonomy organizes approaches into several major branches: Trajectory-Level Synthesis creates new state-action sequences through dynamics models or generative processes; Corrective and Interventional Augmentation incorporates failure recovery and human feedback; Visual and Perceptual Augmentation applies transformations to observation spaces; Invariance and Equivariance-Based methods exploit symmetries; Model-Based Generation uses learned world models; Cross-Domain Transfer leverages data from related tasks; Policy-Guided approaches use learned policies to inform augmentation; Application-Specific methods target domains like robotics or autonomous driving; and Evaluation branches establish theoretical guarantees. Representative works span from trajectory synthesis methods like Mimicgen[1] and DemoGen[5] to corrective approaches such as IntervenGen[7] and visual augmentation techniques like RoCoDA[9].

Within Trajectory-Level Synthesis, a particularly active line explores dynamics-based generation, where learned models produce plausible rollouts to expand limited datasets. Koopman Trajectory Synthesis[0] sits squarely in this branch, using Koopman operator theory to generate trajectories that respect system dynamics. This contrasts with nearby works: DemoGen[5] emphasizes task-conditioned generation for robotic manipulation, while Reverse Augmentation[36] explores backward trajectory construction. The tension between model accuracy and generalization appears across these methods—some prioritize faithful dynamics modeling (Offline Trajectory Optimization[40]), while others focus on diversity and coverage. Koopman Trajectory Synthesis[0] distinguishes itself through its linear operator framework for nonlinear dynamics, offering a middle ground between model fidelity and computational tractability that differs from diffusion-based approaches like those in related generative methods.

Related Works in Same Category

The following **4 sibling papers** share the same taxonomy leaf node with the original paper:

1. Mimicgen: A data generation system for scalable robot learning using human demonstrations

Authors: Mandlekar, Ajay, Nasiriany, Soroush, Ajay Mandlekar, et al. (24 authors total) | **Year/Venue**: 2023 | **URL**: View paper

Abstract

Imitation learning from a large set of human demonstrations has proved to be an effective paradigm for building capable robot agents. However, the demonstrations can be extremely costly and time-consuming to collect. We introduce MimicGen, a system for automatically synthesizing large-scale, rich datasets from only a small number of human demonstrations by adapting them to new contexts. We use MimicGen to generate over 50K demonstrations across 18 tasks with diverse scene configurations, object ...

Relationship Analysis

Both papers belong to the Dynamics-Based Trajectory Generation category, leveraging learned models to synthesize trajectories for offline imitation learning. While the original paper (KATS) uses Koopman operator theory to generate entire trajectories in a lifted linear space with state-equivariant representations, MimicGen adapts human demonstrations by spatially transforming object-centric trajectory segments and executing them with robot controllers. The key difference is that KATS focuses on learning linear dynamics representations in latent space for trajectory synthesis, whereas MimicGen directly transforms and replays demonstration segments in the original state space using object pose transformations.

---

2. DemoGen: Synthetic Demonstration Generation for Data-Efficient Visuomotor Policy Learning

Authors: Xue, Zhengrong, Deng, Shuying, Zhengrong Xue, et al. (16 authors total) | **Year/Venue**: 2025 • Robotics | **URL**: View paper

Abstract

Visuomotor policies have shown great promise in robotic manipulation but often require substantial amounts of human-collected data for effective performance. A key reason underlying the data demands is their limited spatial generalization capability, which necessitates extensive data collection across different object configurations. In this work, we present DemoGen, a low-cost, fully synthetic approach for automatic demonstration generation. Using only one human-collected demonstration per task...

Relationship Analysis

Both papers belong to the Dynamics-Based Trajectory Generation category, leveraging learned models to synthesize trajectories for data augmentation in offline imitation learning. While KATS uses Koopman operator theory to generate entire trajectories in a linearized latent space with policy-equivariant transformations, DemoGen takes a different approach by spatially adapting a single human demonstration to novel object configurations through 3D point cloud editing and scene rearrangement. The key distinction is that KATS focuses on trajectory-level synthesis through dynamics linearization and symmetry transformations, whereas DemoGen emphasizes spatial generalization through geometric manipulation of visual observations.

---

3. Offline Imitation Learning with Model-based Reverse Augmentation

Authors: Jie-Jing Shao, Hao-Sen Shi, Jiejing Shao, Lan-Zhe Guo, Haoran Shi, et al. (7 authors total) | **Year/Venue**: 2024 • Knowledge Discovery and Data Mining | **URL**: View paper

Abstract

In offline Imitation Learning (IL), one of the main challenges is the covariate shift between the expert observations and the actual distribution encountered by the agent, because it is difficult to determine what action an agent should take when outside the state distribution of the expert demonstrations. Recently, the model-free solutions introduced supplementary data and identified the latent expert-similar samples to augment the reliable samples during learning. Model-based solutions build f...

Relationship Analysis

Both papers belong to the Dynamics-Based Trajectory Generation category, leveraging learned models to synthesize trajectories for offline imitation learning. They overlap in addressing covariate shift through trajectory-level augmentation and dynamics modeling. However, the original paper (KATS) uses Koopman operator theory to linearize dynamics in a lifted space with state-equivariant representations, while the candidate paper (SRA) employs reverse dynamics models that generate trajectories backward from expert-observed states in a self-paced manner, fundamentally differing in their augmentation mechanisms and theoretical foundations.

---

4. Offline Trajectory Optimization for Offline Reinforcement Learning

Authors: ZIQI ZHAO, Zhaochun Ren, Liu Yang, Yongli Liang, Fajie Yuan, et al. (10 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

Offline reinforcement learning (RL) aims to learn policies without online explorations. To enlarge the training data, model-based offline RL learns a dynamics model which is utilized as a virtual environment to generate simulation data and enhance policy learning.

However, existing data augmentation methods for offline RL suffer from (i) trivial improvement from short-horizon simulation; and (ii) the lack of evaluation and correction for generated data, leading to low-qualified augmentation. I...

Relationship Analysis

Both papers belong to the Dynamics-Based Trajectory Generation category, leveraging learned models to synthesize trajectories for offline imitation learning. While the original paper (KATS) uses Koopman operator theory to model closed-loop expert policy dynamics and generate trajectories through symmetry transformations in a linearized latent space, the candidate paper (OTTO) employs an ensemble of Transformers (World Transformers) to model open-loop environment dynamics and generates long-horizon trajectories through action perturbation with uncertainty-based correction. The key distinction lies in their modeling approach: KATS focuses on policy-equivariant Koopman representations for behavioral consistency, whereas OTTO emphasizes long-horizon rollouts with explicit uncertainty quantification and correction mechanisms.

Contributions Analysis

Overall novelty summary. The paper proposes KATS, a framework for generating complete multi-step trajectories in offline imitation learning using Koopman operator theory with state-equivariant assumptions and refined generator matrices. It resides in the 'Dynamics-Based Trajectory Generation' leaf, which contains five papers total including the original work. This leaf sits within the broader 'Trajectory-Level Synthesis and Adaptation' branch, indicating a moderately populated research direction focused on generating full trajectories rather than single-step augmentations. The taxonomy shows this is an active but not overcrowded area, with sibling leaves exploring diffusion-based synthesis and demonstration adaptation as alternative trajectory-level approaches.

The taxonomy reveals several neighboring research directions that contextualize this work. The sibling 'Diffusion-Based Trajectory Synthesis' leaf contains four papers using generative models for trajectory creation, representing an alternative paradigm to dynamics-based methods. Adjacent branches include 'Corrective and Interventional Augmentation' (addressing distribution shift through corrective labels) and 'Model-Based Data Generation' (using world models for synthesis). The scope note for the original leaf explicitly excludes diffusion and stitching methods, positioning KATS within dynamics-model approaches that preserve system constraints. This placement suggests the work bridges classical control theory (Koopman operators) with modern imitation learning, occupying a distinct methodological niche.

Among 25 candidates examined across three contributions, the trajectory-level synthesis contribution shows one refutable candidate from 10 examined, while the state-equivariant representation (0 from 5) and refined generator matrix (0 from 10) appear more novel within this limited search scope. The single refutable case for trajectory synthesis suggests some prior work addresses multi-step generation, though the specific combination of Koopman theory with state-equivariance and error-correction mechanisms may differentiate KATS. The computational efficiency and robustness contributions show no clear refutations among their examined candidates, indicating these technical innovations may represent more distinctive advances within the constrained literature sample.

Based on this limited analysis of 25 semantically similar papers, KATS appears to occupy a specialized position combining established Koopman theory with novel efficiency and robustness mechanisms for trajectory synthesis. The search scope covers top semantic matches but cannot claim exhaustiveness across the broader offline IL literature. The taxonomy structure suggests moderate competition in dynamics-based trajectory generation, with the work's novelty likely residing in its specific technical approach rather than the high-level goal of multi-step synthesis.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Trajectory-level synthesis process avoiding compounding errors

Description: The authors introduce a method that generates entire expert trajectories as the base unit for data augmentation, rather than single-step transitions. This approach mitigates the compounding errors common in state-space rollouts and ensures generated trajectories are dynamically consistent within the linear Koopman space.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. High-Quality Trajectory Generation via Domain-Knowledge Enhanced GANs

URL: View paper

Brief Assessment

Domain-Knowledge GANs[66] focuses on human mobility trajectory generation for location-based applications, not offline imitation learning. The candidate addresses error accumulation in auto-regressive mobility models, while the original paper addresses compounding errors in state-space rollouts for RL policy learning—fundamentally different problem domains and technical approaches.

---

2. Time-series generation by contrastive imitation

URL: View paper

Prior Art Analysis

Contrastive Imitation[72] demonstrates prior work on trajectory-level generation that addresses compounding errors. The candidate paper explicitly discusses the problem of compounding errors in autoregressive models and proposes a trajectory-level approach that generates entire trajectories rather than single-step transitions. The paper states that 'autoregressive models trained by mle allow learning and computing explicit transition distributions, but suffer from compounding error during rollouts' and proposes a method that 'directly model the distribution of trajectories p(x1,...,x_t)' to alleviate exposure bias and compounding errors. This shows that trajectory-level synthesis to avoid compounding errors was explored before the original paper's submission.

Evidence

Evidence 1 - **Rationale**: Both papers identify compounding error as a fundamental problem in sequential generation. The candidate explicitly discusses this issue and proposes trajectory-level modeling as a solution. - **Original**: we introduce a trajectory-level augmentation method that uses entire expert demonstrations as its base unit. this approach mitigates the compounding errors common in statespace rollouts and ensures generated trajectories are dynamically consistent within the linear koopman space. - **Candidate**: on one hand, autoregressive models trained by mle allow learning and computing explicit transition distributions, but suffer from compounding error during rollouts. on the other hand, adversarial models based on gan training alleviate such exposure bias, but transitions are implicit and hard to asse...

Evidence 2 - **Rationale**: Both papers articulate the same fundamental challenge: single-step approaches risk compounding errors, while trajectory-level approaches are needed to maintain consistency across multi-step sequences. - **Original**: the success of data augmentation hinges on generating high-quality, dynamically consistent data kolev et al. (2024). common single-step techniques like mixup ignore temporal context, risking the creation of physically unrealistic states that can bias the model. synthesizing entire trajectories is a ... - **Candidate**: owing to the fact that time-series features are generated sequentially, generative modeling in the temporal setting faces a two-pronged challenge: first, a good generator should accurately capture the conditional dynamics of stepwise transitions p(xt|x1,...,x t-1); this is important, as the faithful...

Evidence 3 - **Rationale**: The candidate provides formal theoretical analysis (Lemma 1) showing that local moment-matching leads to $O(t^2\varepsilon)$ error accumulation, directly addressing the compounding error problem that the original paper claims to solve. - **Original**: this reveals the central need in offline il for a framework that can generate novel, consistent trajectories without long-horizon error accumulation. - **Candidate**: lemma 1 let maxf∈rhxx (e h~µs x~πs(·|h) f(h,x) -e h~µs x~πθ(·|h) f(h,x)) ≤ϵ. then Δ ¯fs(θ) ∈o(t2ϵ). proof. appendix a. □ this reveals the problem with modeling conditionals per se: not all mistakes are equal. an objective like equation 2 penalizes unrealistic transitions(h,x) by treating all condi...

Evidence 4 - **Rationale**: The candidate demonstrates (Lemma 2) that trajectory-level moment-matching reduces error to O(tε) instead of $O(t^2ε)$, providing both theoretical justification and a practical method for trajectory-level generation that avoids compounding errors—the exact contribution claimed by the original paper. - **Original**: instead of traditional step-by-step rollouts, kats generates new data by modeling entire expert trajectories in a learned latent space. - **Candidate**: lemma 2 let maxf∈rhxx (e h~µs x~πs(·|h) f(h,x) -e h~µθ x~πθ(·|h) f(h,x)) ≤ϵ. then Δ ¯fs(θ) ∈o(tϵ). proof. appendix a. □ this illustrates why even transition-centric adversarial models such as [12,21] have shown promise in generating realistic trajectories [23-26]. first, unlike trajectory-centric ...

3. Rethinking imitation-based planners for autonomous driving

URL: View paper

Brief Assessment

Rethinking Imitation Planners[67] focuses on autonomous driving with perturbation-based augmentation techniques for single-step transitions, not trajectory-level synthesis in a Koopman framework. The candidate addresses compounding errors through data augmentation strategies like state perturbation and normalization, but does not propose generating entire trajectories as base units in a lifted linear space.

4. Improving Vehicle Trajectory Prediction with Online Learning

URL: View paper

Brief Assessment

Online Learning[71] focuses on online updating of vehicle trajectory predictors using sequential historical data at corner cases, not on trajectory-level synthesis for data augmentation that avoids compounding errors in offline imitation learning.

5. Pseudo-simulation for autonomous driving

URL: View paper

Brief Assessment

Pseudo-simulation[69] focuses on autonomous driving evaluation using pre-rendered synthetic observations from 3D Gaussian splatting, not on trajectory-level data augmentation for offline imitation learning. The candidate addresses evaluation methodology rather than training data synthesis to avoid compounding errors in RL frameworks.

6. Minimizing the accumulated trajectory error to improve dataset distillation

URL: View paper

Brief Assessment

Trajectory Error[68] focuses on dataset distillation for imitation learning, addressing accumulated trajectory errors in synthetic dataset generation. The original paper addresses trajectory synthesis for offline imitation learning using Koopman operators. These are fundamentally different problem domains with distinct technical approaches.

7. Stable Video Infinity: Infinite-Length Video Generation with Error Recycling

URL: View paper

Brief Assessment

Stable Video Infinity[70] addresses error accumulation in video generation through error-recycling fine-tuning for diffusion models, not trajectory-level synthesis for imitation learning data augmentation.

8. CCIL: Continuity-based Data Augmentation for Corrective Imitation Learning

URL: View paper

Brief Assessment

CCIL[2] focuses on generating corrective labels in locally continuous regions using dynamics models for single-step corrections, not trajectory-level synthesis. The original paper's approach of generating entire expert trajectories as base units in Koopman space is technically distinct from CCIL[2]'s neighborhood-based corrective labeling method.

9. Scenediffuser: Efficient and controllable driving simulation initialization and rollout

URL: View paper

Brief Assessment

Scenediffuser[74] focuses on autonomous driving simulation with amortized diffusion for closed-loop rollout, not on offline imitation learning data augmentation. The technical contexts and problem domains are fundamentally different.

10. Policy-guided diffusion

URL: View paper

Brief Assessment

Policy-guided Diffusion[73] generates entire trajectories using diffusion models to avoid compounding errors in autoregressive rollouts. However, it focuses on offline RL with policy guidance for distribution shift, not on Koopman-based trajectory synthesis for imitation learning as in the original paper.

Contribution 2: State-equivariant Koopman representation for computational efficiency

Description: The framework leverages a state-equivariant assumption instead of action-equivariant modeling, which avoids severe computational and memory costs of prior approaches. This design makes KATS highly efficient and scalable for complex tasks by learning only a single operator rather than per-action operators.

This contribution was assessed against **5 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Equivariance and partial observations in Koopman operator theory for partial differential equations

URL: View paper

Brief Assessment

Equivariance Partial Observations[61] focuses on exploiting symmetries in PDEs and partial observation scenarios for Koopman operators, not on computational efficiency trade-offs between state-equivariant versus action-equivariant modeling in offline imitation learning contexts.

2. Dynamics harmonic analysis of robotic systems: Application in data-driven koopman modelling

URL: View paper

Brief Assessment

Dynamics Harmonic Analysis[62] focuses on exploiting morphological symmetries in robotic systems through isotypic decomposition and harmonic analysis, not on state-equivariant versus action-equivariant modeling trade-offs for computational efficiency in offline imitation learning contexts.

3. Koopman operator and its approximations for systems with symmetries

URL: View paper

Brief Assessment

Koopman Symmetries[63] focuses on theoretical properties of Koopman operators with symmetries and computational methods (EDMD, kernel DMD) for general dynamical systems analysis. It does not address state-equivariant versus action-equivariant modeling trade-offs in reinforcement learning contexts or computational efficiency comparisons for RL tasks.

4. KEEC: Koopman Embedded Equivariant Control

URL: View paper

Brief Assessment

KEEC[65] focuses on embedding vector fields and preserving control effects through isometry, not on avoiding per-action operators for computational efficiency. The candidate's state-equivariant approach serves a different purpose than KATS's efficiency-focused design.

5. Koopman-Equivariant Gaussian Processes

URL: View paper

Brief Assessment

Koopman-Equivariant GPs[64] focuses on Gaussian process models with Koopman-equivariance for forecasting and uncertainty quantification in dynamical systems, not on offline imitation learning or trajectory synthesis. The candidate's state-space symmetries serve a different purpose than KATS's state-equivariant design for avoiding per-action operators in RL data augmentation.

Contribution 3: Refined generator matrix to counteract approximation errors

Description: The authors design an adaptive symmetric generator matrix that makes the model more robust to the inherent approximation errors of finite-dimensional Koopman representations. This is achieved through an optimization process weighted by the Koopman model's prediction error, improving the quality of synthesized trajectories.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Identification of Nonlinear Systems Using the Infinitesimal Generator of the Koopman Semigroupâ A Numerical Implementation of the Mauroyâ Goncalves Method

URL: View paper

Brief Assessment

Infinitesimal Generator[57] addresses numerical stability issues in computing the Koopman operator's matrix representation and logarithm for the Mauroy-Goncalves method, not the design of adaptive symmetric generator matrices weighted by prediction error for trajectory synthesis in imitation learning.

2. Robust Nonlinear FIR Filtering via Koopman Operator Framework with Combined Bounded-Gaussian Noise Characterization

URL: View paper

Brief Assessment

Robust Nonlinear FIR[56] addresses generator matrices in the context of Koopman operator approximation errors for filtering with bounded-Gaussian noise, not for trajectory synthesis in offline imitation learning. The technical domains and applications are fundamentally different.

3. Modularized data-driven approximation of the Koopman operator and generator

URL: View paper

Brief Assessment

Modularized Koopman[52] focuses on modularized EDMD for interconnected systems using the Koopman generator structure, not on adaptive symmetric generator matrices weighted by prediction error to counteract approximation errors in finite-dimensional representations.

4. Data-driven approximation of Koopman operators and generators: Convergence rates and error bounds

URL: View paper

Brief Assessment

Koopman Convergence Rates[54] focuses on theoretical convergence analysis and error bounds for approximating Koopman operators from data, not on designing adaptive generator matrices for trajectory synthesis in imitation learning contexts.

5. Finite-Data Error Bounds for Koopman-Based Prediction and Control

URL: View paper

Brief Assessment

Finite-Data Error Bounds[55] focuses on probabilistic error bounds for Koopman operator approximation in control systems, not on designing adaptive symmetric generator matrices to counteract approximation errors in trajectory synthesis for offline imitation learning.

6. EDMD-Based Robust Observer Synthesis for Nonlinear Systems

URL: View paper

Brief Assessment

EDMD Robust Observer[53] focuses on observer design for state estimation in nonlinear systems using Koopman operators, not on data augmentation or trajectory synthesis for imitation learning. The technical contexts are fundamentally different.

7. Operator-informed machine learning: Extracting geometry and dynamics from time series data

URL: View paper

Brief Assessment

Operator-informed ML[58] focuses on approximating the generator matrix for continuous-time dynamical systems from time series data, while the original paper addresses discrete-time Koopman operators in offline imitation learning with trajectory synthesis.

8. Koopman Spectral Analysis and System Identification for Stochastic Dynamical Systems via Yosida Approximation of Generators

URL: View paper

Brief Assessment

Yosida Approximation[59] focuses on stochastic dynamical systems and uses resolvent-based approximations for Koopman generators in SDEs, not on counteracting approximation errors in offline imitation learning through adaptive symmetric generator matrices weighted by prediction error.

9. System identification based on sparse approximation of Koopman operator

URL: View paper

Brief Assessment

Sparse Koopman[51] focuses on sparse approximation methods for system identification, not on adaptive symmetric generator matrices weighted by prediction error for trajectory synthesis in imitation learning contexts.

10. Hybrid Koopman-neural network approach for robust parameter estimation and prediction in duffing oscillators

URL: View paper

Brief Assessment

Hybrid Koopman-neural[60] focuses on parameter estimation and prediction in Duffing oscillators using a hybrid approach. The candidate's context mentions generator matrices and numerical errors but does not demonstrate prior work on adaptive symmetric generator matrices weighted by prediction error for trajectory synthesis in imitation learning.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Koopman-Assisted Trajectory Synthesis: A Data Augmentation Framework for Offline Imitation Learning View paper
- [1] Mimicgen: A data generation system for scalable robot learning using human demonstrations View paper
- [2] CCIL: Continuity-based Data Augmentation for Corrective Imitation Learning View paper
- [3] Guided Data Augmentation for Offline Reinforcement Learning and Imitation Learning View paper
- [4] Variable-speed teachingâ playback as real-world data augmentation for imitation learning View paper
- [5] DemoGen: Synthetic Demonstration Generation for Data-Efficient Visuomotor Policy Learning View paper
- [6] Beyond ood state actions: Supported cross-domain offline reinforcement learning View paper
- [7] IntervenGen: Interventional Data Generation for Robust and Data-Efficient Robot Imitation Learning View paper
- [8] Good data is all imitation learning needs View paper
- [9] RoCoDA: Counterfactual Data Augmentation for Data-Efficient Robot Learning from Demonstrations View paper
- [10] DABI: Evaluation of Data Augmentation Methods Using Downsampling in Bilateral Control-Based Imitation Learning with Images View paper
- [11] Disentangling policy from offline task representation learning via adversarial data augmentation View paper
- [12] Data augmentation for efficient learning from parametric experts View paper
- [13] Dataset Distillation for Offline Reinforcement Learning View paper
- [14] Synthetic experience replay View paper
- [15] D-CODA: Diffusion for Coordinated Dual-Arm Data Augmentation View paper
- [16] Improving Generalization in Game Agents with Data Augmentation in Imitation Learning View paper
- [17] An imitation learning method with data augmentation and post processing for planning in autonomous driving View paper
- [18] NeRF in the Palm of Your Hand: Corrective Augmentation for Robotics via Novel-View Synthesis View paper
- [19] Synthetic data generation using imitation training View paper
- [20] Planning in Autonomous Driving Using Imitation Learning With Research on Data Aggregation View paper
- [21] Towards imitation learning to branch for mip: A hybrid reinforcement learning based sample augmentation approach View paper
- [22] Is behavior cloning all you need? understanding horizon in imitation learning View paper
- [23] Learning from Random Demonstrations: Offline Reinforcement Learning with Importance-Sampled Diffusion Models View paper
- [24] Self-evolved Imitation Learning in Simulated World View paper
- [25] DiffStitch: Boosting Offline Reinforcement Learning with Diffusion-based Trajectory Stitching View paper
- [26] Tube-NeRF: Efficient Imitation Learning of Visuomotor Policies From MPC via Tube-Guided Data Augmentation and NeRFs View paper
- [27] Offline behavior distillation View paper
- [28] Offline imitation learning through graph search and retrieval View paper
- [29] RL-Fill: Timing-Aware Fill Insertion Using Reinforcement Learning View paper
- [30] Diffusion model is an effective planner and data synthesizer for multi-task reinforcement learning View paper
- [31] Unsupervised Data Generation for Offline Reinforcement Learning: A Perspective from Model View paper
- [32] Value Function Evaluation with Data Augmentation for Offline Reinforcement Learning View paper
- [33] Identifying expert behavior in offline training datasets improves behavioral cloning of robotic manipulation policies View paper
- [34] Efficient Imitation Learning with Conservative World Models View paper
- [35] Imitation Learning through Image Augmentation Using Enhanced Swin Transformer Model in Remote Sensing View paper
- [36] Offline Imitation Learning with Model-based Reverse Augmentation View paper

- [37] Improving offline reinforcement learning with inaccurate simulators View paper
- [38] Efficient Deep Learning of Robust Policies From MPC Using Imitation and Tube-Guided Data Augmentation View paper
- [39] ROMA: Reverse Model-Based Data Augmentation for Offline Reinforcement Learning View paper
- [40] Offline Trajectory Optimization for Offline Reinforcement Learning View paper
- [41] Offline inverse reinforcement learning View paper
- [42] Equivariant Data Augmentation for Generalization in Offline Reinforcement Learning View paper
- [43] Synthetic Data is Sufficient for Zero-Shot Visual Generalization from Offline Data View paper
- [44] HIPODE: Enhancing Offline Reinforcement Learning with High-Quality Synthetic Data from a Policy-Decoupled Approach View paper
- [45] BiTrajDiff: Bidirectional Trajectory Generation with Diffusion Models for Offline Reinforcement Learning View paper
- [46] S2p: State-conditioned image synthesis for data augmentation in offline reinforcement learning View paper
- [47] Leveraging Haptic Feedback to Improve Data Quality and Quantity for Deep Imitation Learning Models. View paper
- [48] Synthetic Dataset Generation and Learning From Demonstration Applied to Industrial Manipulation View paper
- [49] SEIL: Simulation-augmented Equivariant Imitation Learning View paper
- [50] ROPA: Synthetic Robot Pose Generation for RGB-D Bimanual Data Augmentation View paper
- [51] System identification based on sparse approximation of Koopman operator View paper
- [52] Modularized data-driven approximation of the Koopman operator and generator View paper
- [53] EDMD-Based Robust Observer Synthesis for Nonlinear Systems View paper
- [54] Data-driven approximation of Koopman operators and generators: Convergence rates and error bounds View paper
- [55] Finite-Data Error Bounds for Koopman-Based Prediction and Control View paper
- [56] Robust Nonlinear FIR Filtering via Koopman Operator Framework with Combined Bounded-Gaussian Noise Characterization View paper
- [57] Identification of Nonlinear Systems Using the Infinitesimal Generator of the Koopman Semigroupâ□□A Numerical Implementation of the Mauroyâ□□Goncalves Method View paper
- [58] Operator-informed machine learning: Extracting geometry and dynamics from time series data View paper
- [59] Koopman Spectral Analysis and System Identification for Stochastic Dynamical Systems via Yosida Approximation of Generators View paper
- [60] Hybrid Koopman-neural network approach for robust parameter estimation and prediction in duffing oscillators View paper
- [61] Equivariance and partial observations in Koopman operator theory for partial differential equations View paper
- [62] Dynamics harmonic analysis of robotic systems: Application in data-driven koopman modelling View paper
- [63] Koopman operator and its approximations for systems with symmetries View paper
- [64] Koopman-Equivariant Gaussian Processes View paper
- [65] KEEC: Koopman Embedded Equivariant Control View paper
- [66] High-Quality Trajectory Generation via Domain-Knowledge Enhanced GANs View paper
- [67] Rethinking imitation-based planners for autonomous driving View paper
- [68] Minimizing the accumulated trajectory error to improve dataset distillation View paper
- [69] Pseudo-simulation for autonomous driving View paper
- [70] Stable Video Infinity: Infinite-Length Video Generation with Error Recycling View paper
- [71] Improving Vehicle Trajectory Prediction with Online Learning View paper
- [72] Time-series generation by contrastive imitation View paper
- [73] Policy-guided diffusion View paper
- [74] Scenediffuser: Efficient and controllable driving simulation initialization and rollout View paper