

Continual world model with counterfactual simulated interactions

Anonymous authors

September 7, 2025

Abstract

This paper proposes a continual world model that improves robustness by prospectively simulating the consequences of a model’s own predictions and rehearsing on those counterfactual trajectories. The setting is bionic limb behavior modeling, where a predictor f maps recent residual-body sensor histories to desired prosthetic joint profiles across multiple locomotion tasks (e.g., level walking, stairs, ramps). The core contribution *multitask prospective rehearsal (MPR)* couples f with a learned prospective dynamics model $g(x_k, y_k) \approx x_{k+1}$ that forecasts the next sensor state given the current state and an action (predicted joint profile). By rolling g forward on f ’s own outputs to generate counterfactual inputs $\hat{x}_{k+1} = g(x_k, \hat{y}_k)$, and pairing these with the future target y_{k+1} , the method creates error-aware rehearsal data that explicitly conditions f to correct the distribution shift it induces. This enables continual, multitask adaptation without online environment interaction or access to expert labels at deployment.

Methodologically, MPR integrates three components. (1) **Multitask decomposition:** a shared temporal backbone f_s is followed by lightweight task-specific heads f_t , balancing transfer of common structure with per-task specialization. (2) **Continual adaptation with evolving heads and rehearsal:** when a new task arrives, a new head is added and both f_s and prior heads are updated using a task-balanced reservoir of rehearsal data drawn from previous tasks. Unlike conventional replay, the rehearsal buffer contains both original validation samples and prospective samples synthesized by g from the model’s own rollouts, at a fixed buffer budget to avoid confounding gains with data volume. (3) **Prospective modeling and joint objectives:** training alternates between (a) supervised learning of f on observed and prospective inputs with targets y_k and y_{k+1} , and (b) system identification of g via a one-step prediction loss $\mathcal{L}_g = \sum_k \|x_{k+1} - g(x_k, y_k)\|_2^2$. Algorithmically, for each time step in a held-out validation split, MPR computes $\hat{y}_k = f(x_{k-T:k})$, synthesizes $\hat{x}_{k+1} = g(x_k, \hat{y}_k)$, and appends $(\hat{x}_{k+1}, y_{k+1})$ to the rehearsal set used to update both f_s and all f_t .

A simple theoretical analysis clarifies the need for such counterfactual rehearsal: even if f is Lipschitz, imitation-learning errors can compound exponentially over time because the predictor’s actions alter future inputs. By minimizing the one-step model error and training f on g ’s rollouts, MPR aims to keep the induced input drift within a small constant bound, thereby mitigating compounding error. This prospective, model-based rehearsal plays a role analogous to imagination-augmented planning in model-based RL, but is trained entirely offline and without rewards, crucial for safety-critical human-machine interaction.

Empirically, the framework is model-agnostic and is evaluated across three real-world motion datasets. Beyond improved accuracy and reduced forgetting relative to regularization, architecture, and replay-based baselines, MPR enhances robustness: gains grow with Jensen-Shannon divergence between train and test distributions, and the method confers resilience to FGSM adversarial perturbations and additive noise. A pilot exoskeleton study further indicates improved stability (lower Lyapunov exponents) of model-driven foot kinematics. Collectively, these results support counterfactual, dynamics-aware rehearsal as a principled mechanism for continual learning under covariate shift in safety-critical imitation learning.

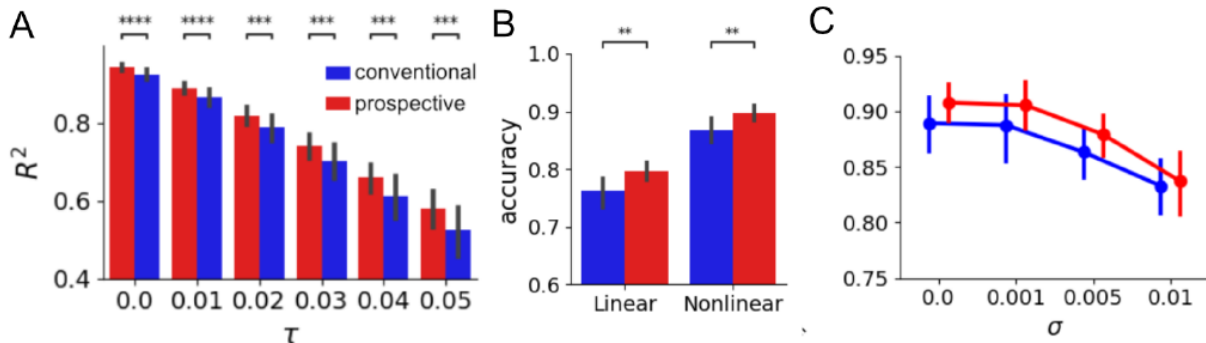


Figure 1: (A) Performance comparison of our counterfactual-interaction-based prospective and conventional rehearsal strategies across different strengths, τ , of adversarial perturbations. (B) Performance of linear and non-linear probing on downstream locomotion task classification. Prospective rehearsal training performs best across the board. (C) Comparison of the performance of continual models with conventional and prospective rehearsal for different levels of input noise. Prospective rehearsal training outperforms the conventional model.