SETMIMIC: POLICY-IN-THE-LOOP INTERPRETATION OF MOTION INTENT FOR HUMANOID IMITATION FROM IMPERFECT MOTION REFERENCES

Anonymous authors
Paper under double-blind review

ABSTRACT

Humanoid imitation commonly relies on high-quality motion capture of professional performers, treating reference joint trajectories as reliable targets for tracking rewards. Internet videos offer a much broader source of motion data, but video-based motion capture often produces physical artifacts such as foot skating, ground penetration, and inconsistent contacts. Directly rewarding joint-level tracking of these references can encourage policies to reproduce artifacts or pursue infeasible targets. Inspired by how professional animators repair motion by preserving the intended action while correcting local artifacts, our key insight is that an imperfect reference can still convey reliable motion intent even when its exact joint trajectories are unsuitable for execution. We present **SetMimic**, a policy-in-the-loop interpretation framework that learns a humanoid policy for each imperfect retargeted motion by preserving reliable intent and resolving uncertain details through robot execution feedback. SetMimic estimates uncertainty to build a structured *Motion Intent Representation*, then samples and refines intent-consistent trajectories into a *reference tube*. Short policy rollouts and a learned feasibility critic guide refinement toward realizations that the robot can execute. An *uncertainty-conditioned set-tracking* objective tracks reliable components precisely while allowing uncertain components to vary within admissible regions, where physical objectives guide execution without a set-tracking penalty. This formulation couples motion interpretation with policy learning to preserve the intended action while accommodating imperfect video-derived references.

1 INTRODUCTION

Humanoid robots promise to acquire broad repertoires of whole-body skills from the rapidly growing supply of human videos and motion-capture data. Recent advances in monocular reconstruction recover increasingly accurate human motion from unconstrained video (Goel et al., 2023; Shin et al., 2024; Zhang et al., 2024b), while learning-based humanoid systems show that large motion collections can support whole-body tracking, teleoperation, and autonomous skill learning (Fu et al., 2024; He et al., 2024a; Ji et al., 2024). A common pipeline reconstructs a human, retargets the reconstruction to a robot embodiment, and trains a physics-based policy to track the result. This pipeline implicitly elevates an estimate to a ground-truth control target, assuming that every reference coordinate is both correct and executable.

That assumption rarely holds for motions recovered from Internet videos and retargeted to robots, where both stages can introduce substantial physical artifacts (Araujo et al., 2025; Yang et al., 2025; Zhang et al., 2026). The resulting references may violate physical constraints, making direct joint-level tracking unreliable. Manual cleanup is costly and does not guarantee executability on the target robot.

Existing pipelines leave a representational gap between motion recovery and policy learning. Stronger kinematic retargeting improves task-space fidelity (Araujo et al., 2025; Zhao et al., 2026), while contact- and dynamics-aware optimization reduces skating, penetration, and implausible interactions (Yang et al., 2025; Zhang et al., 2026). Filters and learned motion priors suppress reconstruction noise (Zhang et al., 2024b). These approaches are valuable upstream, but they typically

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

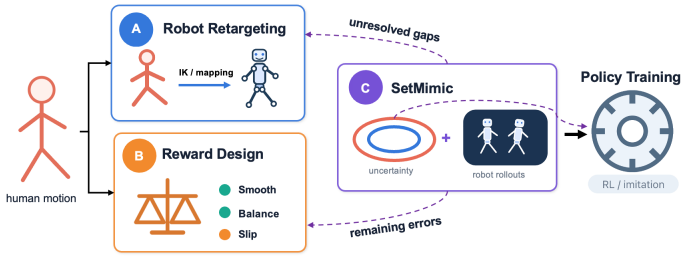


Figure 1: **SetMimic interprets imperfect motion for policy learning.** Much like a human animator who recognizes the intended action before correcting local artifacts, SetMimic interprets the motion conveyed by an imperfect reference and turns it into a precise trajectory that the robot can physically execute.

commit to one nominal trajectory before policy learning. Residual ambiguity is collapsed, and the output is still consumed as a deterministic target.

At the control level, physics-based imitation has advanced from motion-specific policies (Peng et al., 2018) to reusable motion priors and versatile humanoid controllers (Peng et al., 2021; Luo et al., 2023; Tessler et al., 2024; He et al., 2025; Liao et al., 2025). Better exploration, tracking rewards, and physical regularizers can make a controller more tolerant, but they cannot determine whether a particular motion detail reflects the intended action or an upstream artifact. When the target is corrupted, stronger point tracking reproduces the error more faithfully, whereas stronger stability regularization may erase distinctive parts of the action. Reward reweighting merely moves this conflict; it does not remove the false premise that the reference is uniformly trustworthy.

Inspired by professional animators who preserve the intended action while correcting local artifacts, we treat imperfect references as evidence of reliable *motion intent* rather than exact trajectories to reproduce.

We introduce **SetMimic**, a *policy-in-the-loop interpretation* framework that learns one humanoid policy per imperfect retargeted motion using the same algorithm and architecture. Feature- and event-time uncertainty calibrates a structured *Motion Intent Representation* of pose, contact, end-effector, and rhythm cues. Reliable evidence constrains the action, while uncertain details admit alternative realizations.

SetMimic constructs a *reference tube* by sampling uncertain motion components and contact timings, then refining the proposals through constrained inverse kinematics and trajectory optimization. The resulting coherent trajectories and their aligned uncertainty define admissible regions that remain narrow around reliable details and widen where the reference is ambiguous.

Short policy rollouts test candidate executability, and a feasibility critic feeds balance failures, foot slip, and excessive torque back into refinement. The highest-scoring candidate anchors *uncertainty-conditioned set tracking*: reliable components are tracked precisely, while uncertain components incur no tracking penalty within their admissible regions. Unlike reward reweighting around a fixed target, this objective allows balance, actuation, and smoothness to determine the precise execution within those regions. Uncertainty specifies where and how much adaptation is allowed; set tracking lets the controller use that freedom while preserving the action. Figure 1 summarizes this process of preserving reliable intent and resolving uncertain details through execution feedback.

We evaluate SetMimic on 102k frames of highly dynamic motions recovered from internet human videos, where motion-capture and retargeting errors challenge physical execution. We compare against different retargeting methods paired with BeyondMimic and against pretrained generalist trackers. These comparisons assess the benefits of motion intent interpretation for learning faithful and physically credible policies from imperfect motion data.

Our contributions are threefold:

- We formulate imitation from imperfect retargeted motion as **motion-intent interpretation** and introduce a structured Motion Intent Representation whose pose, contact, end-effector, and temporal components are calibrated by feature- and event-time uncertainty.

- We construct a **reference tube** through uncertainty-guided candidate sampling and refinement, and introduce policy-in-the-loop interpretation that uses rollout feedback to retain intent-consistent realizations the target robot can execute.
- We develop an **uncertainty-conditioned set-tracking policy** trained independently for each input motion with the same algorithm and architecture. It combines strict tracking of reliable structure with set-based adaptation in unreliable regions, jointly optimizing motion fidelity and physical executability.

2 Related Work

Human Motion Recovery and Robot Retargeting. Monocular recovery increasingly produces temporally coherent, world-grounded motion (Goel et al., 2023; Ye et al., 2023; Shin et al., 2024; Zhao et al., 2024; Dwivedi et al., 2024). Motion priors improve robustness to incomplete observations (Rempe et al., 2021; Zhang et al., 2024b), and structured 2D priors and penetration-free diffusion improve ambiguous interactions (Han et al., 2026), yet physical artifacts remain. Retargeting bridges human and robot morphology for teleoperation and skill learning (He et al., 2024b; Fu et al., 2024; He et al., 2024a). GMR (Araujo et al., 2025) improves kinematic references; NMR (Zhao et al., 2026) learns a physics-refined robot motion manifold; OmniRetarget (Yang et al., 2025) preserves scene and object interactions; and KDMR (Zhang et al., 2026) enforces dynamics and contact constraints through trajectory optimization. These methods produce a single refined reference, whereas SetMimic retains intent-consistent alternatives for rollout-based controllability assessment.

Physics-Based Motion Imitation for Humanoids. DeepMimic (Peng et al., 2018) learns dynamic skills with tracking rewards, AMP (Peng et al., 2021) introduces adversarial motion priors, PHC (Luo et al., 2023) scales full-body imitation, and MaskedMimic (Tessler et al., 2024) supports partial motion constraints through masked inpainting. Humanoid systems extend these capabilities to teleoperation, expressive tracking, and contact-rich skills (He et al., 2024b;a; Fu et al., 2024; Cheng et al., 2024; Ji et al., 2024; Zhang et al., 2024a; Ma et al., 2026), with advances in sim-to-real transfer, dynamic execution, and generalist tracking (He et al., 2025; Liao et al., 2025; Xie et al., 2025; Luo et al., 2026). BFMTrack (Rupf et al., 2026) uses rollout-based policy gradients to optimize latent sequences for target tracking with a pretrained model. SetMimic instead learns a motion-specific policy with uncertainty-conditioned set tracking, using rollout feedback to identify executable interpretations.

Human Motion Repair. Generative priors support motion synthesis, editing, and goal conditioning (Tevet et al., 2023; Zhang et al., 2023; Dabral et al., 2023; Diomataris et al., 2024). RoHM (Zhang et al., 2024b) reconstructs corrupted motion, PhysDiff (Yuan et al., 2023) adds physics-based projection, and diffusion noise optimization (Karunratanakul et al., 2024) enables denoising and constrained editing. HuTuMotion (Han et al., 2024) and ReinDiffuse (Han et al., 2025) refine motion priors through human feedback and physically motivated reinforcement learning, respectively. Partial motion conditions also guide physics-based control in MaskedMimic (Tessler et al., 2024) and BeyondMimic (Liao et al., 2025). These methods emphasize reconstruction or general physical plausibility; SetMimic addresses constraint uncertainty and robot-specific controllability in retargeted references.

3 Method

3.1 Overview

SetMimic independently learns one reference-conditioned policy per imperfect retargeted motion, using the same architecture and learning algorithm without sharing policy parameters across motions. As shown in Fig. 2, uncertainty estimation identifies reliable evidence to form a Motion Intent Representation; candidate sampling and refinement construct a reference tube of intent-consistent trajectories; policy rollouts provide feasibility feedback to refine these candidates; and uncertainty-conditioned set tracking preserves reliable components while allowing uncertain ones to adapt to robot dynamics.

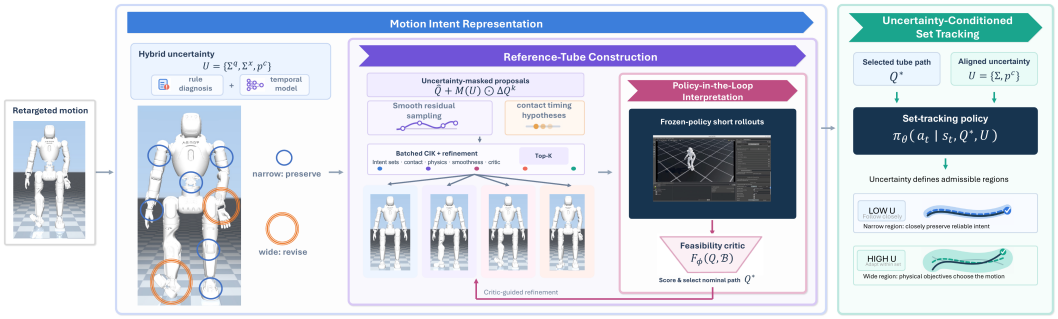


Figure 2: **SetMimic overview.** The figure's four modules correspond to our formulation as follows: uncertainty estimation calibrates the Motion Intent Representation; candidate sampling and refinement instantiate the reference tube; policy-in-the-loop rollouts perform interpretation through execution feedback; and the policy trained for the input motion closely tracks reliable components while set-tracking admissible regions where the input is uncertain.

3.2 Problem Formulation

Let $\hat{\mathbf{Q}}_{1:T} = \{\hat{\mathbf{q}}_t\}_{t=1}^{T}$ denote an imperfect motion retargeted to a robot embodiment $\mathcal{B}$. Each state contains the floating-base pose and joint configuration; optional task-space keypoints and contact cues accompany it as observations. The input may preserve the high-level skill while containing local artifacts caused by perception error, morphology mismatch, inverse-kinematics residuals, or incorrect contact inference.

The practical question is: *which details make this the same action, and which details can change so that the robot can perform it?* For a forward hand sweep, the direction, approximate spatial region, and timing may be clear even when the exact hand coordinates are wrong. Several precise robot trajectories can preserve that intent. We therefore train the policy to balance resemblance to the intended action with physical execution:

$$\max_{\pi} \quad \underbrace{R_{\text{intent}}(\pi)}_{\text{preserve the action}} + \lambda_{\text{phys}} \underbrace{R_{\text{phys}}(\pi)}_{\text{execute it physically}} . \tag{1}$$

Here, the two terms are expected cumulative rewards for intent fidelity and physical behavior. Motion interpretation determines the targets for R_{intent}; balance, contact stability, actuation, and smoothness determine R_{phys}. Although several realizations may be acceptable, each policy rollout produces one precise robot trajectory.

3.3 Motion Intent Representation from Uncertain Motion

SetMimic uses uncertainty to distinguish motion details that should be preserved from those that admit revision, forming a structured Motion Intent Representation. Specifically, the estimated uncertainty identifies imperfect or physically inconsistent segments within the reference motion.

Hybrid uncertainty estimation. We combine normalized kinematic diagnostics, including IK residuals, joint-limit and collision indicators, contact inconsistencies, and velocity or acceleration spikes, with a lightweight temporal model U_ψ trained on paired clean and corrupted motions. Both branches predict per-feature uncertainty and contact probabilities, which are calibrated and averaged element-wise. The fused scores define diagonal covariances $\mathbf{\Sigma}_t^q$ for joint angles and $\mathbf{\Sigma}_t^x$ for end-effector and root pose, together with foot-contact probabilities $\mathbf{p}_t^c$, yielding $\mathbf{u}_t = \{\mathbf{\Sigma}_t^q, \mathbf{\Sigma}_t^x, \mathbf{p}_t^c\}$. Robot-specific executability is assessed separately by the rollout feasibility critic (Section 3.5).

Motion Intent Representation. We represent intent as $\mathcal{M}_t = \{\mathcal{P}_t, \mathcal{C}_t, \mathcal{E}_t, \mathcal{R}_t, \mathcal{U}_t\}$, encoding body pose and relative configuration, contact probabilities and support order, end-effector regions and directions, phase and rhythm, and uncertainty $\mathbf{u}_t$, respectively. For each feature group $\mathbf{z}_{t,g} = f_g(\mathbf{q}_t)$, obtained from the robot configuration or forward kinematics, an admissible region $\mathcal{A}_{t,g}$ surrounds the observation $\hat{\mathbf{z}}_{t,g}$. Its covariance $\mathbf{\Sigma}_{t,g}$ and calibrated threshold δ_g bound the covariance-normalized squared error. Feature differences, denoted $\ominus_g$, use wrapped joint angles

and local rotational coordinates where appropriate. Reliable features yield narrow regions; uncertain ones allow greater variation while preserving motion direction and rhythm. Contact probabilities similarly permit timing adjustments while retaining support order.

3.4 Reference-Tube Construction by Sampling and Refinement

This stage turns the Motion Intent Representation (Section 3.3) into candidate robot trajectories that preserve its reliable cues while varying uncertain details. The resulting *reference tube* replaces a prescribed pose $\mathbf{q}_t = \mathbf{q}_t^{\text{ref}}$ with an admissible set $\mathbf{q}_t \in \mathcal{F}_t$. Complete candidates provide temporally coherent paths through this tube, while aligned uncertainty defines local freedom around each path. Intent consistency defines the tube; policy rollouts assess executability.

Uncertainty-masked proposals. An edit mask $\mathbf{M}(\mathbf{U})$, derived from $\mathbf{U} = \{\mathbf{u}_t\}_{t=1}^{T}$, keeps reliable features nearly fixed and permits larger changes to uncertain ones. We interpolate residuals $\Delta\mathbf{Q}^k$ sampled at sparse temporal control points and propose contact timings that preserve support order, keeping motion and uncertainty annotations aligned. Batched constrained inverse kinematics (CIK) projects these edits to joint space, preserving reliable targets and stance contacts unless the contact hypothesis changes, while enforcing joint limits and self-collision constraints. This yields N coherent candidate trajectories.

Parallel candidate refinement. We refine candidates in parallel using intent-set agreement, contact consistency, physical constraints, temporal smoothness, and predicted feasibility. The intent-set term preserves the admissible regions specified by the Motion Intent Representation; the remaining terms reduce contact errors, constraint violations, and discontinuities. The feasibility term $\mathcal{L}_{\text{feas}}$ is enabled after the critic is initialized (Section 3.5).

Diverse Top-K selection. After discarding invalid candidates, we rank the remainder by the refinement objective and greedily select K paths subject to a minimum pairwise-distance constraint. These paths form $\mathfrak{T}_{\text{ref}} = \{\mathbf{Q}_{1:T}^k\}_{k=1}^{K}$, with aligned uncertainty and contact annotations retaining local admissible variation. They provide concrete realizations of the Motion Intent Representation for rollout evaluation and set tracking.

3.5 Policy-in-the-Loop Interpretation

Short policy rollouts test which intent-consistent candidates the robot can execute, providing controller-specific feedback beyond analytic repair objectives.

Warm-started short rollouts. For each motion, we warm-start a point-tracking policy π_0 and freeze a snapshot within each interpretation round to evaluate all candidates. The same policy is later updated through set tracking. Each candidate is tested over M rollouts of H_r steps with shared initial states and matched perturbation and dynamics-randomization seeds, isolating reference quality from rollout variation.

Controllability scoring. The score $y(\mathbf{Q}^k)$ aggregates tracking accuracy and valid rollout fraction positively, and foot sliding, normalized actuator demand, and joint-limit or collision violations negatively. The tracking term preserves fidelity to the intended action while the remaining terms assess physical execution.

Critic-guided feedback. A feasibility critic $F_\phi(\mathbf{Q}, \mathcal{B})$ regresses rollout scores and ranks candidates. Held fixed during refinement, it supplies $\mathcal{L}_{\text{feas}} = -F_\phi(\mathbf{Q}, \mathcal{B})$ to favor executable references. Subsequent rounds simulate high-scoring, diverse candidates and update the critic with fresh labels as the policy improves. The best candidate provides a coherent nominal path $\mathbf{Q}^*$ with aligned uncertainty and contact annotations for set tracking, retaining the admissible regions around the selected tube branch.

3.6 Uncertainty-Conditioned Set Tracking

Conventional tracking treats every reference target as fully reliable and enforces precise tracking throughout. For imperfect motions, the policy should track reliable components precisely, but allow uncertain components to track a set of targets consistent with the motion intent rather than

a potentially incorrect fixed point. This lets the robot adapt its motion according to balance and dynamics.

From point targets to set targets. For a feature group g, let $\mathbf{z}_{t,g} = f_g(\mathbf{q}_t)$ denote its current value, such as joint angles, root pose, or end-effector position. A point-tracking reward penalizes deviations from $\mathbf{z}^{\text{ref}}_{t,g}$; reducing its weight weakens this penalty but leaves the preferred target unchanged. A set target instead defines an admissible region $\mathcal{S}_{t,g}$ consistent with motion intent, with tracking error

$$d_g(\mathbf{z}_{t,g}, \mathcal{S}_{t,g}) = \min_{\mathbf{y} \in \mathcal{S}_{t,g}} \|\mathbf{z}_{t,g} \ominus_g \mathbf{y}\|_2, \tag{2}$$

where $\ominus_g$ denotes the feature-appropriate difference defined above, including wrapped joint-angle differences and local rotational errors. This distance is zero inside the admissible region, allowing balance, torque, and smoothness objectives to determine the precise realization; outside it, tracking encourages a return. The time-varying regions preserve the action's spatial and temporal structure.

Uncertainty determines the allowed freedom. The selected path $\mathbf{Q}^*$ anchors each target region $\mathcal{S}_{t,g}$ at $\mathbf{z}^{\text{ref}}_{t,g} = f_g(\mathbf{q}^*_t)$, with extent determined by covariance $\mathbf{\Sigma}_{t,g}$ and threshold δ_g. The policy observes proprioception, a reference horizon, uncertainty, and contact probabilities. For $\mathbf{e}_{t,g} = \mathbf{z}_{t,g} \ominus_g \mathbf{z}^{\text{ref}}_{t,g}$, the tracking reward is

$$r^{\text{track}}_t = \sum_g \alpha_g \exp\Big(-k_g \left[\mathbf{e}^\top_{t,g}(\mathbf{\Sigma}_{t,g} + \epsilon \mathbf{I})^{-1}\mathbf{e}_{t,g} - \delta_g\right]_+\Big), \tag{3}$$

where $[x]_+ = \max(0, x)$, $\epsilon > 0$ regularizes the covariance, and $\alpha_g, k_g > 0$ control feature weights and penalty scales. The penalty measures uncertainty-normalized set violations rather than the feature-space distance in Eq. 2. Small covariance enforces precise tracking; larger covariance allows greater variation with no tracking penalty inside the region.

Contact probability $p^c_{t,l}$ blends stance and swing rewards for foot l, while the selected path preserves phase and support order. Balance, torque, smoothness, and joint-limit terms in Eq. 1 guide physical execution within the allowed regions.

4 Experiments

Our experiments ask whether the same learning algorithm can train a physically credible policy from each imperfect retargeted motion. The main comparison separates the effect of upstream retargeting from the effect of downstream policy learning, and tests whether large-scale pretrained motion models can handle the same imperfect inputs. A component ablation examines how motion intent representation, policy-in-the-loop interpretation, and uncertainty-conditioned set tracking contribute to this objective.

4.1 Experimental Setup

Motion data. Our dataset comprises 102k frames from 22 videos of popular routines, each lasting between 30 seconds and 2.5 minutes. We choose highly dynamic motions because rapid whole-body movements and frequent support changes challenge balance and contact consistency, making physical problems such as foot sliding and excessive torque more apparent. We use GVHMR (Shen et al., 2024) as the default video motion-capture model and retarget its estimates with GMR (Araujo et al., 2025) and NMR (Zhao et al., 2026). The resulting references share a 50-fps timeline. We retain motion-capture and retargeting artifacts without additional manual cleanup to evaluate policy learning from imperfect inputs.

Implementation details. We retain the AgiBot X2 embodiment, a 200-Hz simulator, and a 50-Hz policy that outputs target joint positions. The policy observes robot proprioception and a 0.5-s reference horizon with aligned uncertainty and contact probabilities. Candidate construction samples $N = 64$ structured residual trajectories, refines them for 20 iterations, and retains $K = 8$ diverse paths through the reference tube. Candidate evaluation uses 2-s policy rollouts. The uncertainty estimator is a temporal convolutional network with five residual blocks, width 256, kernel size 3, and dilations $\{1, 2, 4, 8, 16\}$. The feasibility critic is a simple multilayer perceptron (MLP) trained to predict candidate executability from short-rollout scores. We optimize the uncertainty estimator

Table 1: Mean performance over 102k motions frames recovered with GVHMR. The retargeting baselines use BeyondMimic with GMR or NMR references. Each task-trained configuration learns one policy per motion. Results are averaged equally across complete rollouts without falls. Clean MPJPE is evaluated on shared reliable reference components after pelvis XY and yaw alignment. All metrics are lower-is-better.

Method	Clean MPJPE (mm) ↓	Foot Sliding (cm/s) ↓	Torque Violation (‰) ↓	Jerk (rad/s^2/step) ↓
Retargeting baselines				
GMR + BeyondMimic	45.984	4.982	0.0539	5.394
NMR + BeyondMimic	44.921	4.612	0.0435	5.289
Generalist trackers				
BFMTrack	45.243	4.531	0.0455	4.932
SONIC	44.993	4.326	0.0419	5.138
SetMimic (Ours)	**44.654**	**3.724**	**0.0364**	**4.425**

and the feasibility critic with Adam at a learning rate of 3×10^{-4}, and train the control policy with PPO at an initial learning rate of 5×10^{-4}. The point-tracking warm start, interpretation rounds, and final set tracking update the same policy for the current motion. We reuse the same architecture and algorithmic configuration across the 22 reference motions. We match the simulation and optimization settings for the task-trained comparisons on each motion, accounting for candidate rollouts in the per-motion simulator-transition budget. Dynamics randomization covers link mass, center of mass, friction, actuator strength, latency, and observation noise.

Baselines. BeyondMimic (Liao et al., 2025) is the primary policy-learning baseline. We train it with GMR and NMR references to compare retargeting methods under a common learning algorithm. For SetMimic and BeyondMimic, we independently train one policy per motion, keeping the architecture and algorithmic configuration fixed across motions within each method. SetMimic uses NMR references in the main comparison.

BFMTrack (Rupf et al., 2026) and SONIC (Luo et al., 2026) are the generalist baselines, both receiving the same GMR references as SetMimic for fair comparison. We retain their pretrained models and prescribed inference procedures, including BFMTrack's per-motion latent sequence optimization. All comparisons require the same robot embodiment, joint convention, simulator, and evaluation conditions; any embodiment-specific adaptations must be documented before reporting results.

Metrics. All metrics are calculated using a fixed control interval Δt. **Clean MPJPE** is the mean Euclidean position error over valid frames on shared reliable reference components. The reference pose is aligned with the robot pelvis XY position and yaw, and the result is reported in millimetres. **Foot Sliding** is the mean horizontal foot velocity while the foot is in contact with the ground. Contact is detected when the vertical ankle contact force exceeds a predefined threshold, after which $\sqrt{v_x^2 + v_y^2}$ is averaged and reported in cm/s. **Torque Violation** is the proportion of valid joint-frame samples for which the magnitude of the requested pre-clamp PD torque, $\tau = K_p(q_{\text{target}} - q) - K_d\dot{q}$, exceeds the configured actuator effort limit, reported in parts per thousand (‰). Motion smoothness is evaluated using **Jerk**, calculated from the third finite difference of joint position over four consecutive frames, divided by Δt^3, and averaged in absolute value over all joints and valid frames. For easier interpretation, the reported value is $\text{Jerk} \times \Delta t$, representing the mean change in joint acceleration between consecutive control steps, in rad/s^2/step. We evaluate each method over 20 rollouts with different random seeds and report the average performance., and no falls were observed across any of the evaluated motions.

4.2 Main Comparison

Quantitative analysis. Table 1 contains three groups. First, GMR + BeyondMimic and NMR + BeyondMimic compare upstream retargeting methods while keeping policy learning fixed. Second, BFMTrack and SONIC evaluate generalist tracking on the NMR references. Finally, SetMimic

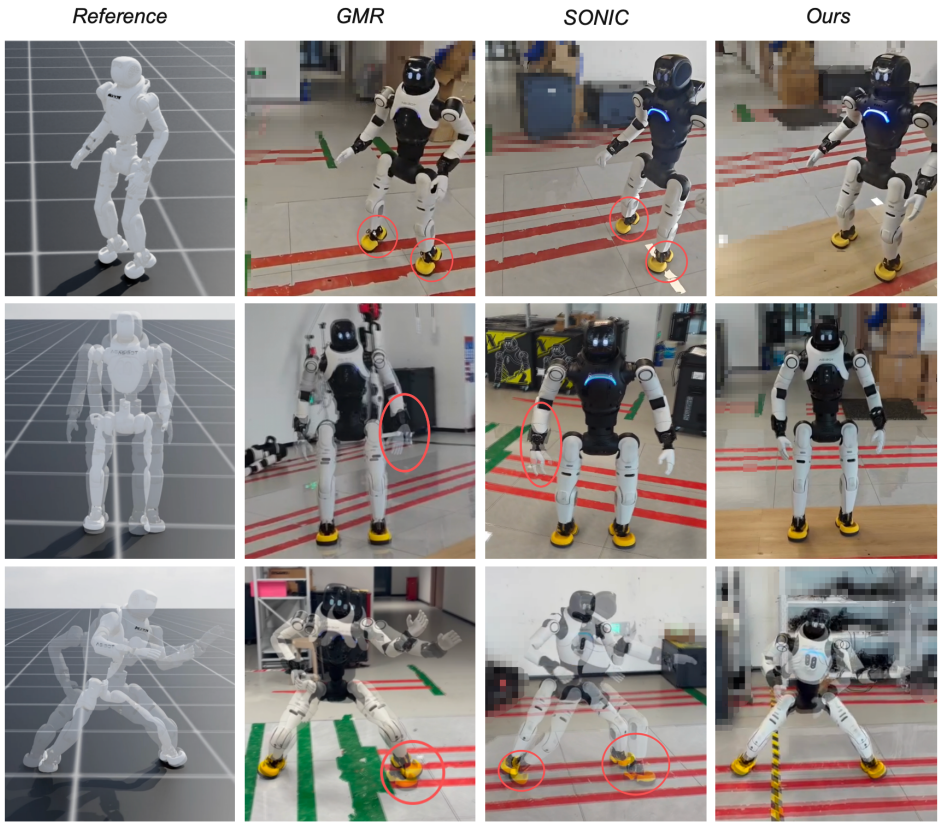


Figure 3: Qualitative real-robot comparison on our dataset. Columns show the imperfect input reference. Each image in the second and third columns overlays two frames to visualize motion jitter and foot stumbling; red circles highlight the affected regions. SetMimic refines imperfect motion references into more executable real-robot motions while retaining the intended movement. See the supplementary videos for a more direct comparison of motion smoothness and foot stability during execution.

uses those same GMR references to learn one policy per motion. Table 1 presents quantitative results over 102k motion frames, showing that SetMimic achieves the lowest error across all four metrics. Among the retargeting baselines, replacing GMR with NMR improves all metrics under the same BeyondMimic tracker, indicating that reference quality contributes to downstream tracking performance. Nevertheless, SetMimic provides further improvements over NMR + BeyondMimic, reducing foot sliding, torque violations, and jerk by 19.3%, 16.3%, and 16.3%, respectively. Since these methods share the same NMR references, these gains demonstrate the benefit of our approach beyond improvements in retargeting alone.

SetMimic also outperforms both generalist trackers. Compared with the strongest baseline for each metric, it reduces foot sliding by 13.9%, torque violations by 13.1%, and jerk by 10.3%, indicating improved contact consistency, better adherence to torque limits, and smoother motion execution. Meanwhile, SetMimic achieves the lowest clean MPJPE of 44.654 mm, although the margin over the best baseline is modest. This pattern is consistent with our method's targeted treatment of imperfect motion components: it improves physical plausibility and temporal smoothness while preserving tracking accuracy on reliable reference components. Together, these results suggest that selectively accommodating reference imperfections enables higher-quality motion execution without sacrificing fidelity to clean motion.

Qualitative analysis. Figure 3 illustrates how reference imperfections manifest during physical execution. The highlighted feet in Taiji and Liancheng show unstable placement, while the overlapping hand positions in Yongchun expose motion jitter in the baseline executions. SetMimic produces

Table 2: Component ablations using GMR as the fixed retargeter, with one independently trained policy per motion. Every variant uses the same source robot motions. PLI denotes policy-in-the-loop interpretation. All metrics are lower-is-better.

Variant	Clean MPJPE (mm) ↓	Foot Sliding (cm/s) ↓	Torque Violation (‰) ↓	Jerk (rad/s^2/step) ↓
Full SetMimic	**44.654**	**3.724**	**0.0364**	**4.425**
Rule-only uncertainty	44.861	4.174	0.0425	5.072
Model-only uncertainty	44.752	4.034	0.0413	4.938
Single candidate ($K = 1$)	44.656	4.317	0.0447	5.322
Without PLI	44.723	4.187	0.0425	5.122
Fixed-width set tracking	44.694	4.096	0.0391	5.217

more stable realizations of these movements without discarding their characteristic poses. These examples qualitatively support interpreting imperfect references through robot execution feedback: allowing adjustments to unreliable motion components can improve physical executability while preserving motion intent. They complement, rather than replace, the aggregate evaluation of tracking fidelity and physical quality.

4.3 Ablation Study

Table 2 presents component ablations using a fixed GMR retargeter and identical source robot motions. The full SetMimic achieves the best performance across all metrics. Using either rule-based or model-based uncertainty alone increases foot sliding, torque violations, and jerk, demonstrating the importance of both sources. Although model-only uncertainty performs better than rule-only uncertainty, their combination provides further gains, supporting their complementary roles in handling imperfect references. Restricting the method to a single candidate ($K = 1$) causes the largest degradation in these three metrics, highlighting the importance of maintaining multiple plausible motion hypotheses.

Removing policy-in-the-loop interpretation consistently degrades motion quality, suggesting that policy feedback helps select reference interpretations compatible with executable behavior. Fixed-width set tracking also worsens performance, particularly jerk, supporting the value of adapting the tracking-set width to reference uncertainty. Meanwhile, clean MPJPE remains nearly unchanged across variants, ranging from 44.654 to 44.861 mm. This is consistent with our design, which selectively acts on imperfect motion components while preserving reliable reference information. Together, these results indicate that the proposed components improve physical plausibility and temporal smoothness without compromising tracking accuracy on clean motion.

5 Conclusion and Limitations

We presented SetMimic, a policy-in-the-loop interpretation framework that independently trains one humanoid policy per imperfect retargeted motion using the same algorithm and architecture. SetMimic uses uncertainty to construct a structured Motion Intent Representation, instantiates it as a reference tube of precise candidate realizations, and uses policy rollouts to identify those the target robot and controller can execute. Each uncertainty-conditioned set-tracking policy follows reliable structure precisely while resolving ambiguous components through robot dynamics and physical objectives. The evaluation averages results from complete rollouts without falls equally across 22 dynamic videos, testing aligned tracking fidelity through Clean MPJPE, and physical execution quality through Foot Sliding, Torque Violation, and Jerk. This formulation connects motion repair with policy learning through an explicit representation of intent and admissible variation.

The current evaluation covers one robot embodiment and a collection of motions with foot contact on the ground. Broader motion repertoires, object and scene contacts, and multi-agent interaction require further evaluation and richer scene modeling. Finally, simulator-based feasibility cannot eliminate sim-to-real mismatch or guarantee hardware safety.

AI USE STATEMENT

We used generative AI tools primarily to polish the writing and improve the clarity, grammar, and readability of the manuscript. All AI-assisted revisions were reviewed and edited by the authors to ensure that they accurately reflected the intended technical meaning. The authors take full responsibility for the final content of this work, including all text, claims, and results.

ETHICS STATEMENT

We have not identified any specific ethical concerns arising from this work.

REPRODUCIBILITY STATEMENT

To facilitate reproducibility, we provide detailed descriptions of the dataset construction procedure and the computation of the evaluation metrics in the paper and accompanying materials. These descriptions document the steps used to prepare the data and evaluate the results, supporting the replication of our experimental evaluation.

REFERENCES

Joao Pedro Araujo, Yanjie Ze, Pei Xu, Jiajun Wu, and C. Karen Liu. Retargeting matters: General motion retargeting for humanoid motion tracking. *arXiv preprint arXiv:2510.02252*, 2025.

Xuxin Cheng, Yandong Ji, Junming Chen, Ruihan Yang, Ge Yang, and Xiaolong Wang. Expressive whole-body control for humanoid robots. *arXiv preprint arXiv:2402.16796*, 2024.

Rishabh Dabral, Muhammad Hamza Mughal, Vladislav Golyanik, and Christian Theobalt. MoFusion: A framework for denoising-diffusion-based motion synthesis. In *CVPR*, pp. 9760–9770, 2023.

Markos Diomataris, Nikos Athanasiou, Omid Taheri, Xi Wang, Otmar Hilliges, and Michael J. Black. WANDR: Intention-guided human motion generation. In *CVPR*, pp. 927–936, 2024.

Sai Kumar Dwivedi, Yu Sun, Priyanka Patel, Yao Feng, and Michael J. Black. TokenHMR: Advancing human mesh recovery with a tokenized pose representation. In *CVPR*, pp. 1323–1333, 2024.

Zipeng Fu, Qingqing Zhao, Qi Wu, Gordon Wetzstein, and Chelsea Finn. HumanPlus: Humanoid shadowing and imitation from humans. *arXiv preprint arXiv:2406.10454*, 2024.

Shubham Goel, Georgios Pavlakos, Jathushan Rajasegaran, Angjoo Kanazawa, and Jitendra Malik. Humans in 4d: Reconstructing and tracking humans with transformers. In *ICCV*, pp. 14783–14794, 2023.

Gaoge Han, Shaoli Huang, Mingming Gong, and Jinglei Tang. HuTuMotion: Human-tuned navigation of latent motion diffusion models with minimal feedback. In *AAAI*, volume 38, pp. 2031–2039, 2024.

Gaoge Han, Mingjiang Liang, Jinglei Tang, Yongkang Cheng, Wei Liu, and Shaoli Huang. ReinDiffuse: Crafting physically plausible motions with reinforced diffusion model. In *IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2218–2227, 2025.

Gaoge Han, Yongkang Cheng, Zhe Chen, Shaoli Huang, and Tongliang Liu. From 2d alignment to 3d plausibility: Unifying heterogeneous 2d priors and penetration-free diffusion for occlusion-robust two-hand reconstruction. In *CVPR*, pp. 42800–42809, 2026.

Tairan He, Zhengyi Luo, Xialin He, Wenli Xiao, Chong Zhang, Weinan Zhang, Kris Kitani, Changliu Liu, and Guanya Shi. OmniH2O: Universal and dexterous human-to-humanoid whole-body teleoperation and learning. *arXiv preprint arXiv:2406.08858*, 2024a.

Tairan He, Zhengyi Luo, Wenli Xiao, Chong Zhang, Kris Kitani, Changliu Liu, and Guanya Shi. Learning human-to-humanoid real-time whole-body teleoperation. *arXiv preprint arXiv:2403.04436*, 2024b.

Tairan He, Jiawei Gao, Wenli Xiao, Yuanhang Zhang, Zi Wang, Jiashun Wang, Zhengyi Luo, Guanqi He, Nikhil Sobanbabu, Chaoyi Pan, Zeji Yi, Guannan Qu, Kris Kitani, Jessica Hodgins, Linxi Fan, Yuke Zhu, Changliu Liu, and Guanya Shi. ASAP: Aligning simulation and real-world physics for learning agile humanoid whole-body skills. *arXiv preprint arXiv:2502.01143*, 2025.

Mazeyu Ji, Xuanbin Peng, Fangchen Liu, Jialong Li, Ge Yang, Xuxin Cheng, and Xiaolong Wang. ExBody2: Advanced expressive humanoid whole-body control. *arXiv preprint arXiv:2412.13196*, 2024.

Korrawe Karunratanakul, Konpat Preechakul, Emre Aksan, Thabo Beeler, Supasorn Suwajanakorn, and Siyu Tang. Optimizing diffusion noise can serve as universal motion priors. In *CVPR*, pp. 1334–1345, 2024.

Qiayuan Liao, Takara E. Truong, Xiaoyu Huang, Yuman Gao, Guy Tevet, Koushil Sreenath, and C. Karen Liu. BeyondMimic: From motion tracking to versatile humanoid control via guided diffusion. *arXiv preprint arXiv:2508.08241*, 2025.

Zhengyi Luo, Jinkun Cao, Kris Kitani, Weipeng Xu, et al. Perpetual humanoid control for real-time simulated avatars. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10895–10904, 2023.

Zhengyi Luo, Ye Yuan, Tingwu Wang, Chenran Li, Fernando Castañeda, Sirui Chen, Zi-Ang Cao, Jiefeng Li, David Minor, Qingwei Ben, Jinhyung Park, David Sami, Zi Wang, Xingye Da, Runyu Ding, Cyrus Hogg, Lina Song, Edy Lim, Eugene Jeong, Tairan He, Haoru Xue, Wenli Xiao, Simon Yuen, Jan Kautz, Yan Chang, Umar Iqbal, Linxi Fan, and Yuke Zhu. SONIC: Supersizing motion tracking for natural humanoid whole-body control. *Science Robotics*, 11(117):eaed4592, 2026. doi: 10.1126/scirobotics.aed4592. URL https://www.science.org/doi/10.1126/scirobotics.aed4592.

Yubiao Ma, Han Yu, Jiayin Xie, Changtai Lv, Qiang Luo, Chi Zhang, Yunpeng Yin, Boyang Xing, Xuemei Ren, and Dongdong Zheng. Robust and generalized humanoid motion tracking. *arXiv preprint arXiv:2601.23080*, 2026.

Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. DeepMimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM TOG*, 37(4):143:1–143:14, 2018.

Xue Bin Peng, Ze Ma, Pieter Abbeel, Sergey Levine, and Angjoo Kanazawa. AMP: Adversarial motion priors for stylized physics-based character control. *ACM TOG*, 40(4):144:1–144:20, 2021.

Davis Rempe, Tolga Birdal, Aaron Hertzmann, Jimei Yang, Srinath Sridhar, and Leonidas J. Guibas. HuMoR: 3d human motion model for robust pose estimation. In *ICCV*, pp. 11488–11499, 2021.

Thomas Rupf, Agon Serifi, David Müller, Sammy Christen, Ruben Grandia, Espen Knoop, and Moritz Bächer. BFMTrack: Latent sequence optimization for physics-based motion tracking with behavioral foundation models. *arXiv preprint arXiv:2606.25056*, 2026.

Zehong Shen, Huaijin Pi, Yan Xia, Zhi Cen, Sida Peng, Zechen Hu, Hujun Bao, Ruizhen Hu, and Xiaowei Zhou. World-grounded human motion recovery via gravity-view coordinates. In *SIGGRAPH Asia Conference Proceedings*, 2024.

Soyong Shin, Juyong Kim, Eni Halilaj, and Michael J. Black. WHAM: Reconstructing world-grounded humans with accurate 3d motion. In *CVPR*, pp. 2070–2080, 2024.

Chen Tessler, Yunrong Guo, Ofir Nabati, Gal Chechik, and Xue Bin Peng. MaskedMimic: Unified physics-based character control through masked motion inpainting. *arXiv preprint arXiv:2409.14393*, 2024.

Guy Tevet, Sigal Raab, Brian Gordon, Yonatan Shafir, Daniel Cohen-Or, and Amit H. Bermano. Human motion diffusion model. In *ICLR*, 2023.

Weiji Xie, Jinrui Han, Jiakun Zheng, Huanyu Li, Xinzhe Liu, Jiyuan Shi, Weinan Zhang, Chenjia Bai, and Xuelong Li. KungfuBot: Physics-based humanoid whole-body control for learning highly-dynamic skills. *arXiv preprint arXiv:2506.12851*, 2025.

Lujie Yang, Xiaoyu Huang, Zhen Wu, Angjoo Kanazawa, Pieter Abbeel, Carmelo Sferrazza, C. Karen Liu, Rocky Duan, and Guanya Shi. OmniRetarget: Interaction-preserving data generation for humanoid whole-body loco-manipulation and scene interaction. *arXiv preprint arXiv:2509.26633*, 2025.

Vickie Ye, Georgios Pavlakos, Jitendra Malik, and Angjoo Kanazawa. Decoupling human and camera motion from videos in the wild. In *CVPR*, 2023.

Ye Yuan, Jiaming Song, Umar Iqbal, Arash Vahdat, and Jan Kautz. PhysDiff: Physics-guided human motion diffusion model. In *ICCV*, pp. 16010–16021, 2023.

Chong Zhang, Wenli Xiao, Tairan He, and Guanya Shi. WoCoCo: Learning whole-body humanoid control with sequential contacts. *arXiv preprint arXiv:2406.06005*, 2024a.

Jianrong Zhang, Yangsong Zhang, Xiaodong Cun, Yong Zhang, Hongwei Zhao, Hongtao Lu, Xi Shen, and Ying Shan. Generating human motion from textual descriptions with discrete representations. In *CVPR*, pp. 14730–14740, 2023.

Siwei Zhang, Bharat Lal Bhatnagar, Yuanlu Xu, Alexander Winkler, Petr Kadlecek, Siyu Tang, and Federica Bogo. RoHM: Robust human motion reconstruction via diffusion. In *CVPR*, pp. 14606–14617, 2024b.

Xiaoyu Zhang, Steven Haener, Varun Madabushi, and Maegan Tucker. Kinodynamic motion retargeting for humanoid locomotion via multi-contact whole-body trajectory optimization. *arXiv preprint arXiv:2603.09956*, 2026.

Qingrui Zhao, Kaiyue Yang, Xiyu Wang, Shiqi Zhao, Yi Lu, Xinfang Zhang, Qiu Shen, Xiao-Xiao Long, and Xun Cao. Make tracking easy: Neural motion retargeting for humanoid whole-body control. *arXiv preprint arXiv:2603.22201v3*, 2026.

Yizhou Zhao, Tuanfeng Y. Wang, Bhiksha Raj, Min Xu, Jimei Yang, and Chun-Hao Paul Huang. Synergistic global-space camera and human reconstruction from videos. In *CVPR*, pp. 1216–1226, 2024.