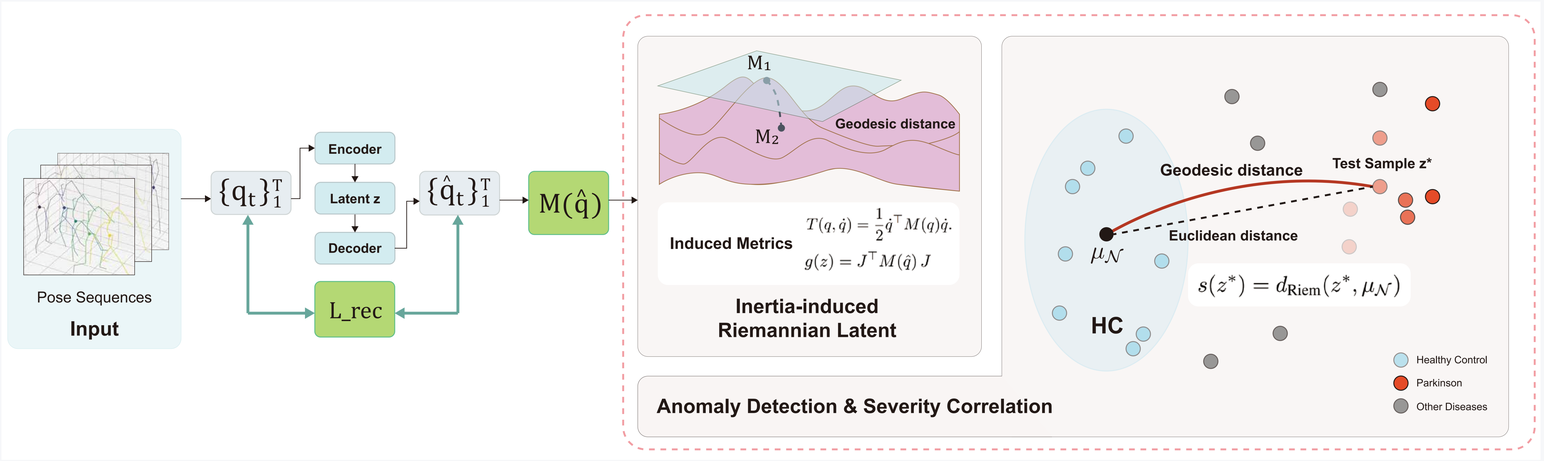
REINS: Learning Inertia-Induced Geometry for Physics-Consistent Motion Representation in Clinical Gait Phenotyping

Yiran Ding[1] Yufei Zhang[2] Zijun Cui[1]

[1]Department of Computer Science and Engineering, Michigan State University, East Lansing, Michigan, USA [2]Independant Researcher

ECCV 2026 Workshop on Human Motion Challenges in Real-World and Clinical Settings (MoCha) · Non-Proceedings Track, Extended Abstract

Overview



Pose sequences are encoded to a latent z; reconstruction is supervised under the inertia metric $M(\hat{q})$ at the decoder output. The latent inherits the pullback metric $g(z) = J^\top M(\hat{q})J$, and each subject is scored by geodesic distance to the healthy-control (HC) reference.

Problem

Gait phenotyping for Parkinson's disease (PD) relies on latent motion representations learned from data alone. Clinical cohorts are small, heterogeneous, and collected under site-specific protocols, so such representations overfit and transfer poorly [1]. Existing physics-informed motion models apply mechanical structure as a **trajectory-level penalty** on reconstruction [5], leaving the **geometry of the representation space** untouched.

Key idea: put physics in the metric

Human gait is the motion of an articulated rigid-body system. Its generalized inertia matrix $M(q)$ is a Riemannian metric on configuration space—the classical kinetic-energy metric [3], applied to the whole moving human body by Neilson et al. [4]:

$$T(q, \dot{q}) = \tfrac{1}{2}\, \dot{q}^\top M(q)\, \dot{q}.$$

$M(q)$ is fixed by mechanics and body shape, so it is available as an inductive bias no matter how little clinical data exists.

Method

Backbone. Encoder–decoder and per-body-part mass/inertia model from PhysPT [5].

Inertia-consistent reconstruction. Velocity error measured in the inertia metric, in place of a trajectory-level Euler–Lagrange residual:

$$\mathcal{L}_{\text{rec}} = \sum_t (\dot{q}_t - \dot{\hat{q}}_t)^\top M(q_t)\, (\dot{q}_t - \dot{\hat{q}}_t).$$

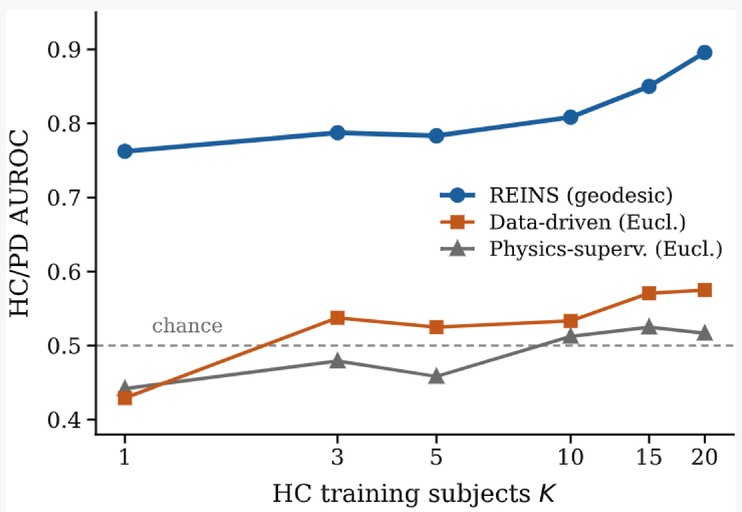
Geodesic read-out. The latent inherits the pullback metric $g(z) = J^\top M(\hat{q})J$ through the decoder Jacobian J [2]. A subject is scored by geodesic distance to the HC reference, approximated on a k-NN graph with edges weighted by local Riemannian length and Dijkstra shortest paths.

Protocol. Fine-tune on CARE-PD [1] DNE healthy controls only. PD and other-movement-disease subjects are never seen in training; clinical labels are used only for evaluation. Baselines share the identical backbone and differ only in latent geometry and read-out: *Data-driven Euclidean* (reconstruction only) and *Physics-supervised Euclidean* (full trajectory-level physics objective), both read out by Mahalanobis distance.

Results

	AUROC ↑		Spearman ρ ↑			
Method	HC/PD	HC/all	PD-GaM	BMCLab	T-SDU-PD	3DGait
Euclidean latent, Euclidean read-out						
Data-driven	0.575	0.609	0.280	0.220	0.257	0.214
Physics-supervised	0.517	0.538	0.251	0.290	0.234	0.170
M-induced Riemannian latent, geodesic read-out						
REINS (ours)	**0.896**	**0.868**	**0.538**	**0.523**	**0.402**	**0.306**
Chance	0.500	0.500	0.000	0.000	0.000	0.000

Left: subject-level anomaly detection on CARE-PD DNE at K=20 HC training subjects (HC/all = HC vs. PD+other movement disease). Right: zero-shot UPDRS-gait severity correlation on four unseen cohorts (n=30, 23, 14, 43). One backbone for all rows.



The physical prior does not need data. HC/PD AUROC as the HC training cohort shrinks (nested subsets).

What the results say

- **Trajectory-level physics supervision does not separate pathological gait.** Both Euclidean read-outs sit near chance on HC vs. PD, including the one trained with a complete trajectory-level physics objective. Re-equipping the *same* encoder with a geodesic distance under M recovers most of the gap.
- **The advantage does not depend on data volume.** The separation survives as the HC training cohort shrinks from 20 subjects to 1.
- **Zero-shot transfer.** The same score correlates with UPDRS-gait on four external cohorts spanning different sites, protocols, and severity distributions, with no fine-tuning, label exposure, or per-cohort calibration.
- **Mechanism.** Adding $\mathcal{L}_{\text{rec}}$ cuts predicted-vs.-ground-truth kinetic-energy error to 22.7%, against 82.1% (reconstruction-only) and 193.4% (full trajectory-level physics objective), at matched pose accuracy (9.79 vs. 9.05 mm test MPJPE).

Conclusion

For clinical motion analysis, physics belongs in the **metric**, not in the loss penalty. Future work broadens the reference cohort across age, body size, and recording protocol.

References

1. Adeli, V., et al.: CARE-PD: A multi-site anonymized clinical dataset for Parkinson's disease gait assessment. NeurIPS Datasets & Benchmarks (2025)
2. Arvanitidis, G., Hansen, L.K., Hauberg, S.: Latent space oddity: on the curvature of deep generative models. ICLR (2018)
3. Bullo, F., Lewis, A.D.: Geometric Control of Mechanical Systems. Springer (2005)
4. Neilson, P.D., Neilson, M.D., Bye, R.T.: A Riemannian geometry theory of human movement: the geodesic synergy hypothesis. Human Movement Science 44, 42–72 (2015)
5. Zhang, Y., Kephart, J.O., Cui, Z., Ji, Q.: PhysPT: Physics-aware pretrained transformer for estimating human dynamics from monocular videos. CVPR (2024)