

Title. *GAITGen: Disentangled Motion-Pathology Impaired Gait Generative Model*

Motivation. Parkinson’s Disease (PD) affects walking mechanics. Machine learning for PD gait analysis is constrained by scarce, imbalanced, and demographically skewed clinical data. Models trained on large, publicly available datasets of healthy individuals fail to capture pathology-specific characteristics that clinicians rely on for accurate assessment. We address this gap by generating clinically relevant gait sequences conditioned on pathology severity (scored by the UPDRS-gait subscore [1]), with controls that separate disease factors from core locomotion dynamics.

Method. We propose GAITGen, a two-stage generative framework that learns a disentangled motion–pathology representation and then synthesizes sequences conditioned on a target severity.

1. **Disentangled Conditional RVQ-VAE.** Two encoders map an input gait sequence to latents: a motion latent z_m intended to be pathology-invariant, and a pathology latent z_p intended to capture severity-specific deviations. Each latent is discretized with Residual Vector Quantization (RVQ) across multiple codebooks to capture hierarchical, physically plausible patterns. Reconstruction uses an SO(3) rotation-aware geodesic loss for rotations plus a L_1 positional term. Disentanglement is enforced with a classifier on z_p , an adversarial classifier with gradient reversal on z_m , capacity asymmetry that constrains z_p , and z_p dropout on healthy sequences.
2. **Hierarchical Token Generation.** We train a bidirectional Mask Transformer with random masking to predict RVQ first-layer tokens conditioned on target severity. A special token separates motion and pathology token streams, which improves cross-stream attention. We then train a Residual Transformer that predicts layer-wise residual tokens, refining fine-grained details across RVQ layers. During inference, masked base tokens are iteratively filled, then residual tokens are rolled out across layers, and the decoder reconstructs the final mesh sequence.
3. **Augmentation in Latent-space.** Disentanglement enables Mix-and-Match composition, where pathology tokens from one sequence are combined with motion tokens from another. This novel augmentation expands underrepresented severe classes without manual labelling.

Dataset. We introduce PD-GaM, a public SMPL [2] 3D mesh dataset of 1,701 walking segments from 30 individuals with PD with UPDRS-gait scores (0 to 3, higher means more severe impairment).

Results. GAITGen outperforms strong motion-generation baselines retrained on PD-GaM. On generation, AVE 0.194 vs 0.898 (MoMask [3]) and 1.037 (MMM [4]). On downstream UPDRS-gait classification, adding GAITGen synthetic data improves F1 score from 0.66 to 0.74 across a participant-held-out test set. Gains replicate across multiple classifiers, including a transformer-based model and a PD-specific GCN. A clinician study with Parkinson’s experts shows near-chance real-vs-synthetic discrimination (precision 0.52, recall 0.57), almost perfect inter-rater agreement on severity scoring (ICC 0.92), and strong alignment between requested severity and clinician-assigned scores for synthetic motions, which supports the clinical face validity of the generated motions.

Ablations. Adversarial suppression of pathology in z_m , capacity reduction in z_p , healthy-only z_p dropout, and rotation-aware loss each contribute to a higher Disentanglement Score and lower reconstruction error. Removing RVQ or removing explicit severity conditioning degrades both disentanglement and fidelity.

Conclusion. GAITGen delivers controlled, pathology-conditioned gait synthesis with explicit disentanglement of motion and pathology, thereby improving clinical realism and downstream severity estimation, and providing a principled path to alleviating data scarcity at higher severities. The approach is model-agnostic at the classifier stage, supports latent-space composition for targeted augmentation, and suggests a viable route toward fairer and more reliable clinical gait analytics.

[1] Goetz, Christopher G., et al. "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results." *Movement disorders: official journal of the Movement Disorder Society* 23.15 (2008): 2129-2170.

[2] Loper, Matthew, et al. "SMPL: A skinned multi-person linear model." *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*. 2023. 851-866.

[3] Guo, Chuan, et al. "Momask: Generative masked modeling of 3d human motions." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.

[4] Pinyoanuntapong, Ekkasit, et al. "Mmm: Generative masked motion model." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.