FoGMap: Wearable Freezing of Gait Detection, Pose Estimation and Mapping in Daily Living Spaces (Extended Abstract)

Ryosuke Hori[1], Taiga Karino[1], Daisuke Ichimura[1], Hiroyuki Deguchi[1,2], Hideo Saito[2], Tsubasa Maruyama[1], and Mitsunori Tada[1]

[1] National Institute of Advanced Industrial Science and Technology
[2] Keio University

Abstract. Freezing of Gait (FoG) is an episodic symptom of Parkinson's disease whose clinical value as a monitoring target depends on capturing when and where it occurs in daily life. We present FoGMap, built around a single mandatory neck-worn camera-and-IMU device that can be augmented with optional wrist and foot IMUs. A modality-flexible multi-task network encodes each sensor and the camera's SLAM-derived neck pose as one token, fuses them by presence-masked attention, and jointly outputs frame-wise FoG probability and 3D full-body pose, the latter via a diffusion decoder. Egocentric localization from the same camera places detected events on a floor plan. On a joint dataset curated from the real WearGait-PD and the synthetic CARE-PD, the neck device alone already detects FoG well, adding sensors improves detection further, and the diffusion decoder preserves gait motion when sensors are missing.

Keywords: Freezing of Gait · Wearable sensing · Human pose estimation · Multi-task learning · Diffusion models

1 Introduction

Freezing of Gait (FoG) in Parkinson's disease (PD) is a paroxysmal interruption of walking and a leading cause of falls [12]. Because it is triggered by turning, doorways, narrow passages, or dual-tasking, brief in-clinic observation cannot capture how it manifests in daily life [9]; the value of FoG monitoring rests on recovering when and where it occurred in everyday activity, the prerequisite for interventions such as environmental modification, cueing, and medication timing.

Wearable FoG detection [3, 11, 14], pose estimation from sparse IMUs [6, 10, 16], and egocentric localization [4] are each mature, but each addresses a single component under a fixed sensor configuration, and FoG detection degrades substantially outside the laboratory [8, 13]. What makes such monitoring deployable is not accuracy alone but what the patient is willing to wear. We therefore make a neck-worn camera-and-IMU device (e.g., THINKLET) the only mandatory element, adding wrist and foot IMUs only when the user accepts them.

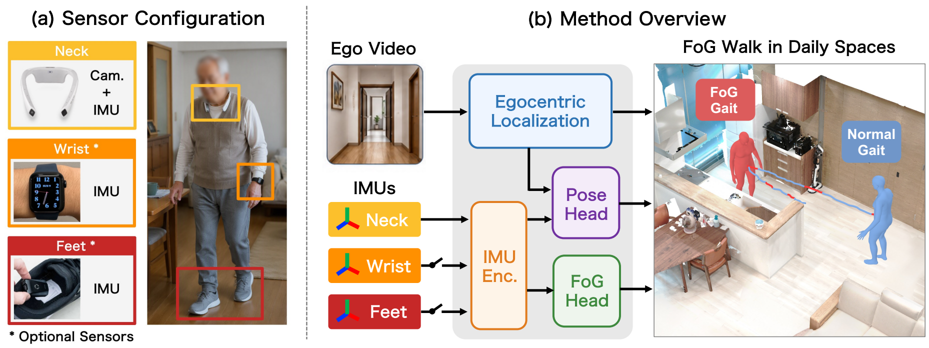

Fig. 1: Overview of FoGMap. **(a)** A mandatory neck-worn camera-and-IMU device, optionally augmented with wrist and foot IMUs, so the wearable burden scales to what the user will wear. **(b)** The IMUs and the SLAM-derived neck pose are fused by presence-masked attention, and a shared backbone branches into a FoG detection head and a diffusion 3D-pose decoder. Egocentric localization from the same camera maps detected events onto the living space.

We propose FoGMap (Fig. 1), which from an arbitrary subset of these sensors jointly estimates frame-wise FoG and 3D full-body pose, and maps detected events onto the living space via egocentric SLAM from the same camera. Pose matters clinically because a binary flag says only that a freeze occurred, whereas the recovered motion shows how it manifests (trembling in place, festination, akinesia). Our contributions are: (1) a curated joint dataset that reconstructs world-frame SMPL for the real WearGait-PD [2] and adds physical-therapist FoG labels to the synthetic CARE-PD [1]; (2) a neck-first, modality-flexible multi-task model that serves any subset of neck, wrist, and foot sensors with one set of weights; (3) its integration with egocentric localization [4] to map FoG events onto a floor plan; and (4) a sensor sweep showing that detection is driven by sensor placement and that a diffusion pose decoder is needed for robustness to missing sensors.

2 Method

FoGMap takes up to five wearable signals as input and treats each as one token, so that any subset of them can be fused by the same network. Two signals come from the mandatory neck-worn device: the neck IMU and the neck pose (6D rotation, height, and velocities, following EgoEgo [7]), which Map-Mono-Ego [4] recovers, together with the camera trajectory, by registering the monocular egocentric video against a pre-acquired point cloud; no raw image enters the network. The remaining three are optional IMUs at the left wrist and both feet. Each available signal is encoded by its own MLP with a learnable token embedding, and presence-masked attention pooling fuses the tokens frame by frame, treating a missing sensor as an absent token rather than as zero padding;

sensors are also dropped at random during training so that the model learns to work with any subset. A bidirectional GRU then turns the fused sequence into a shared feature over 3 s windows at 100 Hz.

Two heads read this feature. Frame-wise FoG detection is well posed, so a small deterministic classifier outputs $P(\mathrm{FoG}_t)$, trained with a focal loss. Pose from a few IMUs is under-determined and a regression head collapses unconstrained joints to a static mean pose, so we solve it generatively with a window-wise bidirectional DiT diffusion decoder [5, 15] that predicts the 23 non-root SMPL joint rotations. At our data scale the decoder trains only when its target is standardized per dimension; SNR loss weighting and ancestral sampling, standard in prior work, inflate high-frequency jitter and are not used. The network is trained end-to-end on the sum of both losses, each averaged over a per-task validity mask so that frames lacking one label still train the other head.

3 Data

WearGait-PD [2] provides real 13-point IMU, plantar-pressure insoles, and frame-wise FoG labels but no pose; we reconstruct world-frame SMPL from its raw, uncalibrated IMU by recovering per-sensor yaw in closed form and snapping the trajectory to the synchronized pressure walkway, yielding pseudo ground truth. CARE-PD [1] provides mocap-derived SMPL, from which we synthesize virtual IMUs, but no FoG labels; a licensed physical therapist annotated FoG intervals from rendered motion. Both use a fixed, FoG-rate-stratified 80/20 split (subject-wise for CARE-PD; trial-wise for WearGait, whose FoG frames are concentrated in two subjects). Models are trained on the union of both training splits.

4 Experiments

Table 1 reports the sensor sweep on the CARE-PD test split (unseen subjects, mocap ground truth, diffusion pose head). We report threshold-free AUROC/AUPRC for FoG, since F1 is prevalence-sensitive, and for pose MPJPE together with the gait-band ratio, the predicted-to-GT amplitude ratio of leg motion in the 0.5–3 Hz band (1.0 ideal; < 1 means the legs do not swing).

Three findings emerge. (1) Detection is driven by sensor placement: each added sensor improves AUROC and AUPRC, and the same trend holds on real WearGait IMU (AUROC .927 $\rightarrow$ .954) and for a regression pose head or a detection-only TCN baseline [8]. The freeze-band (3–8 Hz) power ratio between FoG and non-FoG segments is comparable at the neck (2.09) and the foot (1.95), so the neck device is an informative FoG sensor in its own right. (2) A regression pose head attains a lower MPJPE (62.6 mm with all sensors) but a gait-band ratio of only 0.88; only diffusion keeps leg swing near or above 1.0, which MPJPE alone cannot capture. (3) When the foot IMUs are dropped at test time, diffusion preserves leg swing (0.99 vs. 0.72 for regression), so we adopt it for robustness.

Table 1: Sensor sweep on CARE-PD (diffusion pose head, mixed training, seed means). Gait ratio: 1.0 is ideal. The last row uses the full model but drops the foot IMUs at test time without retraining.

	FoG detection		Pose	
Sensors	AUROC ↑	AUPRC ↑	MPJPE ↓	gait ratio
neck only	.898	.731	81.0 mm	1.05
+ wrist	.921	.779	75.9 mm	1.10
+ wrist + feet	**.943**	**.818**	**75.3 mm**	1.19
+ wrist + feet (feet dropped)	.896	.744	75.0 mm	0.99

For mapping, Map-Mono-Ego on egocentric video rendered from the placed test motions in a scanned environment recovers the trajectory with a median error of 1.1 cm, so the neck pose used as conditioning is recoverable at deployment-relevant accuracy and detected FoG events can be placed on the recovered indoor trajectory (Fig. 1(b)).

5 Conclusion

FoGMap couples a neck-first, modality-flexible multi-task model with egocentric localization to record when, where, and how FoG occurs in daily living spaces, and the sensor sweep gives concrete guidance on the minimum a patient must wear. Limitations include subject leakage in the WearGait split, pseudo-GT pose for WearGait, single-annotator FoG labels for CARE-PD, and localization validated only in a simulated environment. Deploying the intended hardware on PD patients and validating end-to-end on real recordings is our next step.

Acknowledgements. This work was supported by Council for Science, Technology and Innovation, “Cross-ministerial Strategic Innovation Promotion Program (SIP), Development of foundational technologies and rules for expansion of the virtual economy” (JPJ012495). (funding agency: NEDO)

References

1. Adeli, V., Klabučar, I., Rajabi, J., Filtjens, B., Mehraban, S., Wang, D., Seo, H., Hoang, T.H., Do, M.N., Muller, C., de Oliveira, C.N., Coelho, D.B., Ginis, P., Gilat, M., Nieuwboer, A., Spildooren, J., McKay, J.L., Kwon, H., Clifford, G., Esper, C., Factor, S.A., Genias, I., Dadashzadeh, A., Shum, L., Whone, A., Mirmehdi, M., Iaboni, A., Taati, B.: CARE-PD: A multi-site anonymized clinical dataset for Parkinson’s disease gait assessment. In: Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track (2025), `https://arxiv.org/abs/2510.04312`, dataset DOI: 10.5683/SP3/TWIKMK

2. Anderson, A.J., Eguren, D., Gonzalez, M.A., Caiola, M., Khan, N., Watkinson, S., Zuccaroli, I., Hirczy, S.S., Zabetian, C.P., Mills, K., Moukheiber, E., Moro-Velazquez, L., Dehak, N., Motley, C., Muir, B.C., Butala, A., Kontson, K.: WearGait-PD: An open-access wearables dataset for gait in Parkinson's disease and age-matched controls. Scientific Data **13**(1), 440 (2026). https://doi.org/10.1038/s41597-026-06806-2
3. Bächlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J.M., Giladi, N., Tröster, G.: Wearable assistant for Parkinson's disease patients with the freezing of gait symptom. IEEE Transactions on Information Technology in Biomedicine **14**(2), 436–446 (2010). https://doi.org/10.1109/TITB.2009.2036165
4. Deguchi, H., Hori, R., Amaya, K., Maruyama, T., Tada, M., Saito, H.: Map-Mono-Ego: Map-grounded global human pose estimation from monocular egocentric video. In: Proc. IEEE International Conference on Image Processing (ICIP) (2026)
5. Du, Y., Kips, R., Pumarola, A., Starke, S., Thabet, A., Sanakoyeu, A.: Avatars grow legs: Generating smooth human motion from sparse tracking inputs with diffusion model. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 481–490 (2023). https://doi.org/10.1109/CVPR52729.2023.00054
6. Huang, Y., Kaufmann, M., Aksan, E., Black, M.J., Hilliges, O., Pons-Moll, G.: Deep inertial poser: Learning to reconstruct human pose from sparse inertial measurements in real time. ACM Transactions on Graphics **37**(6), 1–15 (2018). https://doi.org/10.1145/3272127.3275108, proc. SIGGRAPH Asia
7. Li, J., Liu, C.K., Wu, J.: Ego-body pose estimation via ego-head pose estimation. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 17142–17151 (2023). https://doi.org/10.1109/CVPR52729.2023.01644
8. Lin, W., Grewal, S.S.: Deep learning for freezing of gait detection: Cross-dataset validation reveals critical deployment gaps between laboratory and daily living wearable monitoring. Sensors **26**(4), 1352 (2026). https://doi.org/10.3390/s26041352
9. Mancini, M., McKay, J.L., Cockx, H., D'Cruz, N., Esper, C.D., Filtjens, B., Heimler, B., MacKinnon, C.D., Palmerini, L., Roerdink, M., Young, W.R., Hausdorff, J.M.: Technology for measuring freezing of gait: Current state of the art and recommendations. Journal of Parkinson's Disease **15**(1), 19–40 (2025). https://doi.org/10.1177/1877718X241301065
10. Mollyn, V., Arakawa, R., Goel, M., Harrison, C., Ahuja, K.: IMUPoser: Full-body pose estimation using IMUs in phones, watches, and earbuds. In: Proc. CHI Conference on Human Factors in Computing Systems. pp. 1–12. ACM (2023). https://doi.org/10.1145/3544548.3581392
11. Moore, S.T., MacDougall, H.G., Ondo, W.G.: Ambulatory monitoring of freezing of gait in Parkinson's disease. Journal of Neuroscience Methods **167**(2), 340–348 (2008). https://doi.org/10.1016/j.jneumeth.2007.08.023
12. Nutt, J.G., Bloem, B.R., Giladi, N., Hallett, M., Horak, F.B., Nieuwboer, A.: Freezing of gait: moving forward on a mysterious clinical phenomenon. The Lancet Neurology **10**(8), 734–744 (2011). https://doi.org/10.1016/S1474-4422(11)70143-0
13. Sigcha, L., Borzì, L., Olmo, G.: Deep learning algorithms for detecting freezing of gait in Parkinson's disease: A cross-dataset study. Expert Systems with Applications **255**, 124522 (2024). https://doi.org/10.1016/j.eswa.2024.124522

14. Sigcha, L., Borzì, L., Pavón, I., Costa, N., Costa, S., Arezes, P., López, J.M., De Arcas, G.: Improvement of performance in freezing of gait detection in Parkinson's disease using Transformer networks and a single waist-worn triaxial accelerometer. Engineering Applications of Artificial Intelligence **116**, 105482 (2022). https://doi.org/10.1016/j.engappai.2022.105482
15. Tevet, G., Raab, S., Gordon, B., Shafir, Y., Cohen-Or, D., Bermano, A.H.: Human motion diffusion model. In: International Conference on Learning Representations (ICLR) (2023)
16. Van Wouwe, T., Lee, S., Falisse, A., Delp, S., Liu, C.K.: DiffusionPoser: Real-time human motion reconstruction from arbitrary sparse sensors using autoregressive diffusion. In: Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2513–2523 (2024). https://doi.org/10.1109/CVPR52733.2024.00243