Lightweight Patient-Invariant Model for Freezing of Gait Detection with Cohort Clustering

Nasimuddin Ahmed
TCS Research & Innovation
Kolkata, India
nasim.ahmed@tcs.com

Aniruddha Sinha
TCS Research & Innovation
Kolkata, India
aniruddha.s@tcs.com

Avik Ghose
TCS Research & Innovation
Kolkata, India
avik.ghose@tcs.com

***Abstract*—Freezing of Gait (FoG) is one of the most debilitating symptoms of Parkinson's Disease (PD), severely impairing mobility and increasing the risk of falls. Traditional FoG detection methods predominantly rely on offline machine learning models, limiting their feasibility for real-time monitoring and wearable deployment. In this study, we present a lightweight and patient-invariant FoG detection framework, specifically optimized for real-time execution on resource-constrained hardware. Our approach utilizes Kullback-Leibler (KL) divergence to measure the similarity between a new patient's gait features and existing cohort data, enabling an adaptive and generalizable detection model. Additionally, we introduce a cohort selection strategy, categorizing subjects into Matched Clusters (MC) and Unmatched Clusters (UC) to enhance model robustness while reducing data dependency. Experimental evaluations on the Daphnet dataset demonstrate that our patient-invariant model achieves comparable classification accuracy, with an average *Sensitivity* of .91 and *Specificity* of .80, while significantly reducing model size. Furthermore, our approach effectively operates with a single IMU sensor, making it a computationally efficient and practical solution for continuous FoG assessment in real-world applications.**

***Index Terms*—Freezing of Gait (FoG), Parkinson's Disease (PD), IMU, Patient-Invariant Model, Real-Time FoG Assessment**

I. Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disorder affecting millions worldwide, leading to severe motor impairments and a decline in overall quality of life. Among its most disabling symptoms is Freezing of Gait (FoG)—a phenomenon where individuals suddenly lose the ability to initiate or continue walking, often resulting in falls and injuries. The unpredictable nature of FoG significantly compromises patient safety and independence, making its accurate assessment and management a critical aspect of PD care. Since there is no cure for PD, treatment primarily focuses on symptomatic management using a combination of pharmacological interventions, physical therapy, and assistive strategies. Levodopa and dopamine agonists are commonly prescribed to alleviate motor symptoms, but their effectiveness fluctuates over time, often leading to motor complications such as dyskinesia and worsening FoG episodes. Given these challenges, continuous monitoring of FoG is crucial for personalized treatment adjustments, optimizing medication schedules, and improving patient outcomes. Currently, FoG assessment is primarily conducted in clinical settings, where patients perform some predefined tasks under medical supervision. While such assessments provide valuable insights, they are inherently limited to controlled environments and fail to capture the full range of symptom variations experienced in daily life. FoG episodes are highly context-dependent, influenced by environmental factors, cognitive load, and emotional states, which cannot be fully replicated in a clinical setting. Furthermore, infrequent clinical visits lead to gaps in symptom tracking, delaying necessary treatment adjustments and potentially increasing the risk of falls. A home-based, real-time FoG detection system would allow for continuous symptom monitoring, providing clinicians with a more comprehensive understanding of FoG progression. Additionally, real-time detection could enable early intervention strategies, preventing severe episodes and improving patient safety. Beyond passive monitoring, real-time FoG detection has the potential to integrate with assistive systems designed to prevent or mitigate FoG episodes.

With this context, wearable devices, particularly those embedded with Inertial Measurement Unit (IMU) sensors, offer a promising avenue for continuous and real-time assessment of FoG in home environments. With their compact and ergonomic design, IMUs enable unobtrusive gait monitoring, facilitating uninterrupted symptom tracking beyond clinical settings. In recent years, extensive research has been conducted to develop FoG detection or prediction methods utilizing machine learning [1], [2], [6] and deep learning [3], [4] techniques, achieving high accuracy in controlled settings. Bachlin et al. [1] pioneered the use of IMU sensors, leveraging the Freezing Index feature for FoG detection. Mazilu et al. [2] introduced a more sophisticated machine learning framework, demonstrating improved detection performance. A notable deep learning approach was proposed by Rubén et al. [3], who implemented a Convolutional Neural Network (CNN) and introduced the Contextual Windows (CW) method, which incorporates temporal dependencies by including features from adjacent time windows. Additionally, Pham et al. [4] investigated an unsupervised anomaly detection approach utilizing an Adaptive Thresholding-based Anomaly Score Detector (ASD), demonstrating its effectiveness in identifying FoG events using IMU sensors. More recently, Ahmed et al. [5]

developed a robust patient-independent model leveraging a single ankle-mounted accelerometer sensor, achieving substantial accuracy for real-time FoG detection.

While these methods demonstrate high classification performance, they primarily operate in an offline setting, making them unsuitable for real-time deployment on wearable hardware. Additionally, most existing approaches rely on conventional machine learning models, raising a critical question: Which model type—patient-dependent, patient-invariant, or hybrid—is best suited for real-world, real-time FoG detection in wearable systems?
To address these limitations, this study introduces a lightweight and deployable FoG detection system that utilizes only a single IMU sensor, specifically designed for real-time execution on wearable devices. The key contributions of this work are as follows:

- **Optimized Deployability:** Development of a computationally efficient model tailored for resource-constrained wearable hardware, ensuring real-time execution while maintaining high detection accuracy.

- **Lightweight Patient-Invariant model construction:** Leveraging pre-existing cohort data to build a generalizable FoG detection model, thereby minimizing the dependency on extensive patient-specific training data. This approach employs Kullback-Leibler (KL) divergence to assess the similarity between the feature distributions of new patients and existing cohorts, ensuring optimal model adaptation for previously unseen individuals.

By integrating these advancements, our proposed methodology bridges the gap between controlled laboratory-based studies and real-world deployment, significantly enhancing the feasibility of continuous and real-time FoG monitoring in everyday settings.

II. Methodology

As previously stated, the central objective of this study is to facilitate the real-world deployment of a FoG detection system. Motivated by this goal, rather than developing the entire algorithm from scratch, we have leveraged key components from our prior research [5] to construct an optimized and deployable system. The complete methodological framework is outlined in the following section.

A. Preprocessing

The preprocessing stage encompasses signal segmentation, annotation, and conditioning. The dataset is divided into fixed time windows for further analysis. We have chosen a window length of 4 seconds, with a 75% overlap to effectively capture transitions between stages. Each signal window is labeled as Freezing of Gait (FoG) if at least 30% of its samples are classified as *FoG*; otherwise, it is categorized as *Motion (M)*.

B. Dataset Description

In this research, we have utilized the Daphnet dataset, a widely recognized benchmark for FoG detection. This dataset comprises accelerometer recordings from three sensor placements: shank (S) (just above the ankle), thigh (T), and lower back (B) of PD patients. It includes 10 PD patients (7 males, 3 females) with a mean age of 66.5 ± 4.8 years. The dataset's acquisition protocol and detailed characteristics have been extensively documented in previous studies, particularly by Bachlin et al. [1].

C. Noise Reduction Using EWT

The gait pattern of PD patients is intricate, non-stationary, and often affected by sporadic noise. To mitigate noise, we employ the Empirical Wavelet Transform (EWT), which adaptively constructs a wavelet filter bank based on the signal's spectral characteristics. The signal is then decomposed into multiple modes, where noisy components are removed, and only the relevant modes are retained to reconstruct a cleaner signal. A comprehensive discussion on EWT-based mode decomposition and signal reconstruction for FoG detection is available in our previous research [1].

D. Feature Extraction

We have examined a range of statistical and physiological features from existing literature. Based on our analysis, we have selected six statistical features— ***Mean, Median, Standard Deviation, Max, Min, and Range*** — along with five physiological features— ***Freezing Index (FI), FoG Power, Motion Power, Mean Frequency, and Median Frequency***— as outlined in our previous study [1] for FoG prediction. In total, 44 features are extracted at each time step from 3D accelerometer data, including the resultant acceleration ($ACC_x, ACC_y, ACC_z, Resultant$).

E. Feature Selection

The Recursive Feature Elimination (RFE) technique is applied to identify the most significant features, enhancing classifier performance and reducing computational complexity. As outlined in Section II-D RFE is used to rank the 44 extracted features. Experimental results indicate that selecting the top 15 features achieves optimal performance. A comprehensive analysis is available in our previous research.

F. Development of a Lightweight Patient-Invariant Model

When a new patient is diagnosed with PD, deploying a wearable-based system for FoG assessment possess a significant challenge due to the dearth of sufficient individualized data. A major hurdle is the difficulty in obtaining a comprehensive dataset containing both motion and FoG gait patterns. While motion data can be recorded in clinical settings, capturing FoG episodes is impractical since they are strongly influenced by environmental factors and the patient's cognitive state. As a result, employing a patient-dependent model for FoG assessment at an early stage is not feasible. Instead, we propose leveraging pre-existing

cohort data to construct a patient-invariant model, which can be gradually refined by incorporating the patient's own data over time.

Cohort Selection Using Similarity-Based Clustering With this research framework, we hypothesize that a personalized FoG detection model can be effectively built by incorporating only those cohort patients whose gait feature distributions closely match the new patient. To achieve this, the probability distributions of extracted features from the new patient's motion data are compared against those of the cohort patients. Only matching cohorts are integrated into the model, while others are discarded. This selective approach is expected to yield a more robust and efficient model for FoG detection and assessment.

To validate our hypothesis, we utilized the Daphnet dataset, as described in Section II-B. We selected only patients whose data contained both Motion and FoG events, and from the 10 available patients, 8(**S01, S02, S03, S05, S06, S07, S08, S09**) are chosen for analysis. Each patient was treated as a new subject, with the assumption that only 10% of their motion data was available for initial model generation. From this limited motion data, features were extracted and their probability distributions were compared with those of the existing dataset using Kullback-Leibler (KL) divergence, which quantifies differences between probability distributions. The objective of this approach is to establish two cohort clusters based on feature similarity. Each feature comparison yields a KL divergence value, which is then normalized between 0 and 1 using an exponential function to generate a Similarity Score (SS), where values closer to 1 indicate a higher similarity in gait patterns. The Mean Similarity Score (MSS) is then computed by averaging all Similarity Scores (SS) for a given cohort.

Based on this similarity assessment, patient cohorts are categorized into two distinct clusters:

1) **Matched Cluster (MC)** – Cohorts whose mean similarity score exceeds an empirically chosen threshold, indicating strong feature similarity with the new patient. Data from these cohorts is incorporated into the patient-invariant model.
2) **Unmatched Cluster (UC)** – Cohorts with a mean similarity score below the threshold, suggesting dissimilar gait patterns compared to the new patient.

A similarity matrix is generated for each patient in tabular form to facilitate cohort selection. An example of the similarity matrix for patient S02 is presented in Table I. In this matrix, each row represents a different patient, while each column corresponds to a specific gait feature and its similarity score. The second-to-last column contains the Mean Similarity Score (MSS), which represents the overall similarity of a cohort's gait pattern to that of the new patient. If the MSS exceeds a predefined threshold, the cohort's data is incorporated into the model, ensuring that only relevant gait patterns contribute to the final FoG detection framework. Figure 1 illustrates the distribution of a selected feature for the new patient compared with a patient from the Matched Cluster (MC) cohort. Conversely, Figure 2 shows the distribution of the same feature for the new patient in comparison with a patient from the Unmatched Cluster (UC) cohort.

Incorporating Data from the Unmatched Cluster for Model Generalization:

While the primary model is constructed using data from the Matched Cluster (MC), incorporating a small fraction of data from the Unmatched Cluster (UC) improves model generalization. Relying exclusively on highly similar data may result in overfitting, limiting the model's adaptability to real-world variations. Introducing a controlled portion of UC data helps mitigate bias and enhances the model's robustness by exposing it to diverse gait characteristics. To optimize performance while maintaining specificity, the proportion of UC data is systematically adjusted. For a clearer understanding of this model generation process, we have outlined the approach in Algorithm 1.

Algorithm 1 Development of a Patient-Invariant Model

0: **Input:** New patient motion data D_n, Cohort dataset D_c, Threshold τ
0: **Output:** Patient-invariant model M
0: **Step 1: Feature Extraction**
0: **for** each patient p_i in D_c **do**
0: Extract gait features F_i from motion data
0: **end for**
0: Extract gait features F_n from new patient data
0: Select 10% of motion data from F_n for similarity computation
0: **Step 2: Compute Similarity Matrix**
0: **for** each cohort patient p_i in D_c **do**
0: **for** each feature f_j in F_n **do**
0: Compute KL divergence: $D_{KL}(F_n^j || F_i^j)$
0: Normalize similarity score: $S_{ij} = \exp(-D_{KL})$
0: **end for**
0: Compute Mean Similarity Score (MSS_i) as average of S_{ij}
0: **end for**
0: **Step 3: Cohort Selection**
0: **for** each cohort patient p_i **do**
0: **if** $\text{MSS}_i \geq \tau$ **then**
0: Assign p_i to Matched Cluster (MC)
0: **else**
0: Assign p_i to Unmatched Cluster (UC)
0: **end if**
0: **end for**
0: **Step 4: Model Construction**
0: Train model M using data from MC
0: Augment with 10% of motion data from each UC patient for generalization
0: **Return** optimized model MOD =0

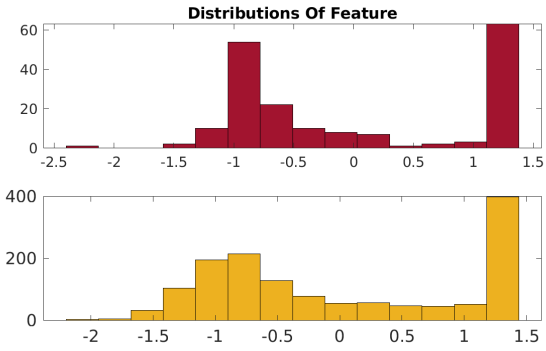

Fig. 1. Distribution of particular feature for both new patient and a patient from MC cohort

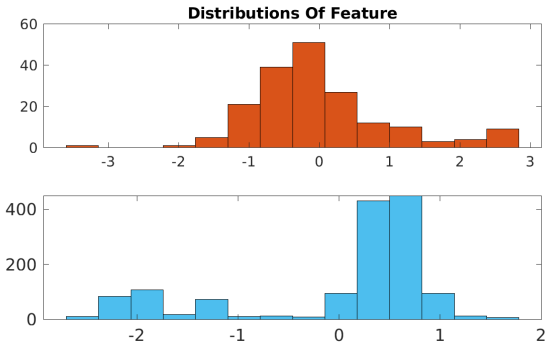

Fig. 2. Distribution of particular feature for both new patient and a patient from UC cohort

G. Classification Module

Due to the infrequent occurrence of FoG, the training data is significantly imbalanced. To address this, we applied random undersampling to the majority class, where surplus Motion (M) data points are randomly removed to achieve a balanced dataset. Although more sophisticated sampling techniques are available, our study does not primarily focus on sampling methods. Random undersampling is chosen for its simplicity, effectiveness, and ability to reduce computational overhead. FoG detection is framed as a binary classification problem, with ***FoG (F)*** as the positive class and ***Motion (M)*** as the negative class. For classification, we employ an ensemble of bagged decision trees as a supervised machine learning approach.

III. Results & Discussion

FoG detection performance is conventionally assessed using ***Sensitivity*** $\left(SN = \frac{TP}{TP+FN}\right)$ and ***Specificity*** $\left(SP = \frac{TN}{TN+FP}\right)$, which are widely adopted in the literature. SN quantifies the proportion of FoG windows correctly identified by the model, whereas SP measures the proportion of Motion windows accurately classified.

Given the focus of this study on real-world deployability, performance evaluation extends beyond classification accuracy to include model efficiency, particularly memory consumption—a critical factor for deployment in resource-constrained environments. We conducted a systematic analysis to examine the trade-off between classification performance and model size as the proportion of data from the Unmatched Cluster (UC) was varied. The results, summarized in the Table III, indicate that while incorporating additional UC data does not yield significant improvements in classification accuracy, it leads to a substantial increase in model size. This expansion may negatively impact the feasibility of deploying the model on low-power embedded systems.
To ensure robustness and generalizability, model performance was evaluated using K=10 cross-validation, where SN and SP values were averaged across ten validation folds. Additionally, when the UC data proportion was set to 100%, the model's performance closely aligned with our previous research findings, demonstrating consistency. Since prior studies do not report model size, this evaluation is primarily benchmarked against our earlier work.

To benchmark our proposed algorithm, we present a comparative analysis against prior studies that utilized the Daphnet dataset [1] with the patient-independent model. Table II summarizes the results, highlighting key performance metrics such as ***SN***, ***SP***, Window Size (WS), and Model Size.

San et al. [3] optimized their approach by selecting an optimal probability threshold to maximize sensitivity, which, however, led to a notable reduction in specificity. To ensure a fair comparison, we applied the same thresholding strategy and reported our results as OurAlgo-O, while OurAlgo-D represents the model's performance with the default classifier threshold. Both models were developed using a patient-invariant approach, incorporating 10% of data from the Unmatched Cluster cohort for enhanced generalization.

Similarly, in our previous research Ahmed et al. [5], we conducted a comparable experiment, reporting the results as Nasim-O for the optimal model and Nasim-D for the default classifier. As our focus is on developing a lightweight, deployable system, we have also included model size as a critical evaluation metric. The significantly reduced model size of 0.451 MB in OurAlgo-D and OurAlgo-O highlights the efficiency of our approach, ensuring comparable sensitivity while achieving better specificity than both San et al. [3] and optimal model(Nasim-O) of our previous work Ahmed et al. [5]. Notably, OurAlgo-O and OurAlgo-D utilize only a single IMU sensor (shank placement above the ankle), yet their accuracy remains comparable to or even surpasses that of other approaches that rely on data from all three sensors. This further validates the effectiveness and feasibility of our algorithm for real-world deployment.

Considering these findings, it is evident that our approach significantly reduces model size while maintaining high classification performance, making it a more practical solution for real-time deployment.

TABLE I
SIMILARITY MATRIX FOR PATIENT S02

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	MSS	Cluster Cohort
S01	0.47	0.23	0.78	0.67	0.90	0.73	0.75	0.87	0.49	0.35	0.68	0.42	0.65	0.90	0.90	**0.63**	MC
S03	0.38	0.19	0.09	0.74	0.77	0.62	0.63	0.46	0.75	0.95	0.58	0.10	0.49	0.03	0.08	0.45	UC
S05	0.05	0.16	0.21	0.81	0.73	0.61	0.88	0.91	0.61	0.30	0.66	0.04	0.61	0.08	0.09	0.45	UC
S06	0.06	0.22	0.13	0.67	0.57	0.82	0.98	0.75	0.49	0.42	0.01	0.14	0.69	0.04	0.34	0.42	UC
S07	0.02	0.35	0.11	0.62	0.14	0.74	0.48	0.68	0.19	0.67	0.12	0.11	0.91	0.07	0.36	0.37	UC
S08	0.59	0.33	0.88	0.75	0.94	0.54	0.79	0.29	0.15	0.89	0.35	0.41	0.38	0.90	0.74	0.61	MC
S09	0.61	0.38	0.84	0.24	0.51	0.64	0.86	0.84	0.49	0.95	0.89	0.43	0.27	0.59	0.65	0.62	MC

TABLE II
COMPARATIVE RESULTS WITH THE STATE OF THE WORK

	SN	SP	WS	Model Size In MB
Bachlin [1]	.73	.82	4	
Mazilu [2]	.66	.95	4	
San [3]	.95	.74	4	
Nasim-O [5]	.95	.70	4	1.3
Nasim-D [5]	.83	.83	4	1.3
OurAlgo-D	.81	.82	4	.451
OurAlgo-O	.91	.80	4	.451

TABLE III
PERFORMANCE METRICS WITH DIFFERENT PERCENTAGE OF UNMATCHED CLUSTER

OtherDataPercentage	SN	SP	SIZE in MB
10	0.81	0.82	.451
30	.82	.82	.902
50	.82	.83	.921
70	.82	.83	1.1
90	.81	.83	1.3
100	.82	.83	1.4

IV. CONCLUSION

This study presents a real-time, patient-invariant FoG detection framework optimized for wearable deployment. By leveraging KL divergence-based cohort selection, the model effectively adapts to new patients while maintaining high detection accuracy and reduced model size for resource-constrained devices. Incorporating only 10% of motion data for similarity-based cohort selection, along with controlled augmentation from the Unmatched Cluster (UC), enhances generalization without compromising efficiency. Future work will explore transfer learning to augment UC data, improving model adaptability and robustness for unseen patients, further enhancing real-world FoG detection.

REFERENCES

[1] M. Bachlin, M. Plotnik, D. Roggen, I. Maidan, J. M. Hausdorff, N. Giladi, and G. Troster, "Wearable assis- tant for parkinson's disease patients with the freezing of gait symptom," IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 2, pp. 436–446, 2010.

[2] S. Mazilu, M. Hardegger, Z. Zhu, D. Roggen, G. Tröster, M. Plotnik, and J. M. Hausdorff, "Online detection of freezing of gait with smartphones and machine learning techniques," in 2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops, 2012, pp. 123–130.

[3] Rubén San-Segundo, Honorio Navarro-Hellın, Roque Torres-Sánchez, Jessica Hodgins, and Fernando De la Torre, "Increasing robustness in the detection of freez- ing of gait in parkinson's disease," Electronics, vol. 8, no. 2, pp. 119, 2019.

[4] T. T. Pham, S. T. Moore, S. J. G. Lewis, D. N. Nguyen, E. Dutkiewicz, A. J. Fuglevand, A. L. McEwan, and P. H. W. Leong, "Freezing of gait detection in parkin- son's disease: A subject-independent detector using anomaly scores," IEEE Transactions on Biomedical En- gineering, vol. 64, no. 11, pp. 2719–2728, 2017.

[5] N. Ahmed, S. Singhal, V. Sharma, S. Bhattacharya, A. Sinha and A. Ghose, "A Patient-Invariant Model for Freezing of Gait Detection Aided by Wavelet Decomposition," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 1015-1019, doi: 10.1109/ICASSP39728.2021.9414780.

[6] N. Ahmed, S. Singhal, A. Sinha and A. Ghose, "A Patient Invariant Model Towards the Prediction of Freezing of Gait," ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10095187.

[7] Parkinson's Foundation May. 2025. [Online]. Available: https://www.parkinson.org/understanding-parkinsons/statistics