# DeepFoG: An IMU-Based Detection of Freezing of Gait Episodes in Parkinson's Disease Patients via Deep Learning

Thomas Bikias[1], Dimitrios Iakovakis[1], Stelios Hadjidimitriou[1], Vasileios Charisis[1] and Leontios J. Hadjileontiadis[1,2]*

[1]Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Greece, [2]Department of Electrical and Computer Engineering, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates

Freezing of Gait (FoG) is a movement disorder that mostly appears in the late stages of Parkinson's Disease (PD). It causes incapability of walking, despite the PD patient's intention, resulting in loss of coordination that increases the risk of falls and injuries and severely affects the PD patient's quality of life. Stress, emotional stimulus, and multitasking have been encountered to be associated with the appearance of FoG episodes, while the patient's functionality and self-confidence are constantly deteriorating. This study suggests a non-invasive method for detecting FoG episodes, by analyzing inertial measurement unit (IMU) data. Specifically, accelerometer and gyroscope data from 11 PD subjects, as captured from a single wrist-worn IMU sensor during continuous walking, are processed via Deep Learning for window-based detection of the FoG events. The proposed approach, namely DeepFoG, was evaluated in a Leave-One-Subject-Out (LOSO) cross-validation (CV) and 10-fold CV fashion schemes against its ability to correctly estimate the existence or not of a FoG episode at each data window. Experimental results have shown that DeepFoG performs satisfactorily, as it achieves 83%/88% and 86%/90% sensitivity/specificity, for LOSO CV and 10-fold CV schemes, respectively. The promising performance of the proposed DeepFoG reveals the potentiality of single-arm IMU-based real-time FoG detection that could guide effective interventions via stimuli, such as rhythmic auditory stimulation (RAS) and hand vibration. In this way, DeepFoG may scaffold the elimination of risk of falls in PD patients, sustaining their quality of life in everyday living activities.

**Keywords: Parkinson's disease, freezing of gait, deepFoG, smartwatch, deep learning, rhythmic auditory stimulation**

## 1 INTRODUCTION

Parkinson's Disease (PD) is a progressive neurological disorder related to multiple motor symptoms which affect patients' movement and stability. One of the most common and most disturbing symptoms is Freezing of Gait (FoG) (Giladi et al., 2001). FoG is characterized by the inability to walk through narrow corridors or to take short and fast steps. This results in difficulty in initiating gait or turning while walking, despite the intention of the patient (Nutt et al., 2011; Heremans et al., 2013). Although FoG episodes usually make their presence in the late stages of PD or other Parkinsonian syndromes, in cases where the motor symptoms are already intense, FoG episodes may also occur in

early stages (Hall et al., 2015). The latter is due to the lack of timely medication. The pathological hallmark of FoG points to disturbances in the frontal cortical regions, the basal ganglia, and the midbrain locomotor region which are the most probable origins (Nutt et al., 2011). FoG has been associated with cerebellum functionality (Fasano et al., 2017) and has been often described as a derangement of the cognitive dynamics (Pozzi et al., 2019). Moreover, occurrence of FoG has been interdependent on the emotional condition of the PD patient, anxiety and/or stress levels (Witt et al., 2019). Despite the increasing research interest in finding the exact causes and characteristics of FoG events, the biological and neurological nature of FoG has not been fully decoded. Importantly, deep understanding of FoG is significantly impeded by the fact that the profile of FoG may differ substantially amongst patients.

Manifestation of FoG episodes harshly downgrades PD patients' Quality-of-Life (QoL) and affects their mental condition and self-esteem. Concretely, PD patients are restricted from performing essential functions, such as walking initiation and direction change. Furthermore, they experience mood alterations which gradually lead to serious mental health issues. Together with emotional swings, the greater risk of living with FoG is the sudden and instantaneous akinesia, which raises the possibility of falls and injuries (Bloem et al., 2004). This problem is intensified in the case of the elderly PD patients, and especially those who are unattended, wherein their bones are more vulnerable, and hence the chances of FoG occurrences are much higher.

The confrontation of FoG events is an essential necessity that concerns the medical society. Multiple solutions have been proposed to eliminate the FoG symptoms. Medication treatment, including use of levodopa (Fietzek et al., 2013) or monoamine oxidase B inhibitor, rasagiline (Coria and Cozar-Santiago, 2008), is commonly provided to deal with the severity of the motor symptoms and, in particular, FoG episodes. Although several potential drug treatments have been suggested (Zhang et al., 2016), pharmaceutical intake frequently fails to alleviate the symptoms, while, in some cases, its effect lasts only during the drug-ON period. Moreover, medication may very often cause side effects (Hauser, 2009) which lead to further inconveniences to patients. Finally, the estimated expenses of medication are rather high and can become a significant financial burden. Invasive solution of Deep brain stimulation has also been introduced in the event of medication failure or in the extreme condition of total inability to walk (Huang et al., 2018). Nevertheless, the neurological surgery is rarely recommended due to the high risk of dangerous or fatal side effects caused by the direct cerebral intervention.

Alternative studies propose improvement of FoG episodes' characteristics through cognitive exercises (Walton et al., 2018). However, given that the PD patients vary in their cultural, educational, and cognitive characteristics, tailored and personalized cognitive exercise should be considered. Exercises and training with rhythmic cue have indicated that FoG episodes can be countered when patients synchronize their movement with external stimulus (Allen et al., 2010). Rhythmic Auditory Stimulation (RAS) constitutes a solution that may be acceptable for the majority of PD patients. RAS offers improvement in gait (Dalla Bella et al., 2017) via synchronization of the patient's steps with an external acoustic cuing that helps to overcome the failure of walking initiation. In addition, it has been observed that PD patients are less likely to encounter FoG episodes when they walk in coordination with stimulus (Thaut et al., 1996; Lim et al., 2005; Hausdorff et al., 2007). Currently, such solutions require the physical presence of physicians and medical staff. However, application of RAS outside medical sites would allow for restoring patients' independence, while would limit the required Healthcare resources. This would be a significant step forward towards the effective facilitation of the increasing number of PD patients (Dorsey et al., 2007). At large, detection of the FoG events is crucial to be performed timely, and, therefore, a real-time RAS process could be valuable for eliminating FoG *per se* and its related risks.

In this view, multiple research groups have investigated Fog detection using methods which rely on sensor placement in different parts of the patient's body. Essentially, these methods can be divided into two categories based on whether a sensor is placed on the upper limb of the patient or not. Jovanov et al. (2009); Niazmand et al. (2011); Bachlin et al. (2009) have suggested the use of multiple devices attached on the lower and middle body to capture FoG events, using accelerometer and gyroscope sensors. Mazilu et al. (2015a) proposed a gait-assist system with sensors worn on both ankles to detect FoG events from four Fast Fourier Transformation (FFT) features by using Machine Learning. Overall, the majority of prior work utilizes measuring devices which are allocated on the ankles, the hips or the waist of the patient (**Table 1**).

Contrary to the aforementioned works, other studies (Ziv et al. (1999); Vercruysse et al. (2012)) have reported a significant correlation between FoG episodes and upper limb motions. In this respect, various methods have been introduced for detecting FoG episodes from sensors placed on the upper limbs. Specifically, Cole et al. (2011) exploited accelerometer sensors attached on the arm, the shank, and the thigh, along with Electromyography sensors (EMG) to compute to detect instances of FoG episodes using Dynamic Neural Networks. Their method achieved 83% sensitivity and 97% specificity. In another study, Tripoliti et al. (2013) have associated FoG events with wrist movement. Despite of the satisfactory accuracy (82%/90% sensitivity/specificity), their method required the use of five additional sensors (worn in ankles, waist, and chest), leading to inconvenience and thus low user acceptance.

Installation of devices in multiple body parts is rather doubtful due to the inconvenience caused by the volume, weight, difficulty in placement of the hardware and/or discomfort in movement. Exclusive association between FoG and wrist movement has been proposed by Mazilu et al. (2015b) who introduced a Machine Learning model to distinguish FoG episodes from other events during walking (i.e., turns, stops, obstacle avoidance). They used time and frequency features from accelerometer and gyroscope data using sensors attached at both wrists. Their classification pipeline achieved 90%/83% sensitivity/specificity.

Consequently, exploiting the increasing rise of smartwatch sensors, the present study investigates the feasibility of a single

**TABLE 1 |** Comparison of the proposed single-wrist approach FoG detection with other studies exploiting one or more sensors to capture movement.

| | Sensors | Sensors placement | Realtime | Results |
|---|---|---|---|---|
| *Without sensor in upper limb* | | | | |
| Bachlin et al. (2009) | 3 | Ankle, thigh and lower back | Yes | 73.1% sensitivity 81.6% specificity |
| Jovanov et al. (2009) | 2 | Both sensors on belt or knee or ankle or shoe | Yes | unknown |
| Niazmand et al. (2011) | 5 | Shank and belt | No | 88.3% sensitivity 85.3% specificity |
| Mazilu et al. (2015a) | 2 | Ankles | Yes | 97% hit rate |
| Palmerini et al. (2017) | 3 | Shins, lower back | Yes | 83% sensitivity 67% specificity |
| Camps et al. (2018) | 1 | Waist | Yes | 92.6% sensitivity 88.7% specificity |
| *With sensor in upper limb* | | | | |
| Cole et al. (2011) | 3 | Shank, thigh and arm | No | 83% sensitivity 97% specificity |
| Tripoliti et al. (2013) | 6 | Wrists, ankles, waist and chest | No | 81.9% sensitivity 98.7% specificity |
| Mazilu et al. (2015b) | 2 | Both wrists | Yes | 90% sensitivity 83% specificity |
| DeepFoG | 1 | Wrist | Yes | 83% sensitivity 88% specificity |

wrist-based inertial measurement unit (IMU) FoG classification, namely, DeepFoG, for facilitating the real-time detection of FoG episodes. The initial hypothesis is that data collected only from one wrist should provide sufficient information for effectively predicting FoG events. We proposed a non-invasive user-centric solution that can be easily applied in everyday life. Given that FoG is a complex activity task and the research knowledge on this domain is still being explored, manual feature engineering would be rather complicated, and thus application of Deep Learning could be suitable for single-wrist detection (Wang et al., 2019). DeepFoG is based on the training of a Deep Learning model that automatically detects FoG events and differentiates them from stops and walking with turns. The ecological validity of DeepFoG aims towards achieving higher user acceptance, while providing comparative detection performance when compared to the prior art.

# 2 MATERIALS AND METHODS

A Deep Learning approach is being proposed for the detection of FoG events from stops and turns during gait. Our proposed framework is based on linear and angular acceleration data from one wrist that are used as input to a Convolutional Neural Network (CNN). The CNN is trained and evaluated using the CuPiD dataset (Mazilu et al., 2012). These data are fed to the CNN classifier to detect FoG episodes (Class 1) from Walking-with-Turns (Class 2) and Stops (Class 3) in a sliding window manner. Firstly, sensitivity and specificity scores from the CNN implementation are calculated and compared with the respective ones from traditional Machine Learning classifiers. Subsequently, the best-performing model is compared to the state-of-the-art methods. The performance of the models is evaluated in two schemes, namely, a 10-Fold Cross-Validation (CV) (10-CV) and a Leave-One-Subject-Out (LOSO) Cross-Validation (LOSO-CV).
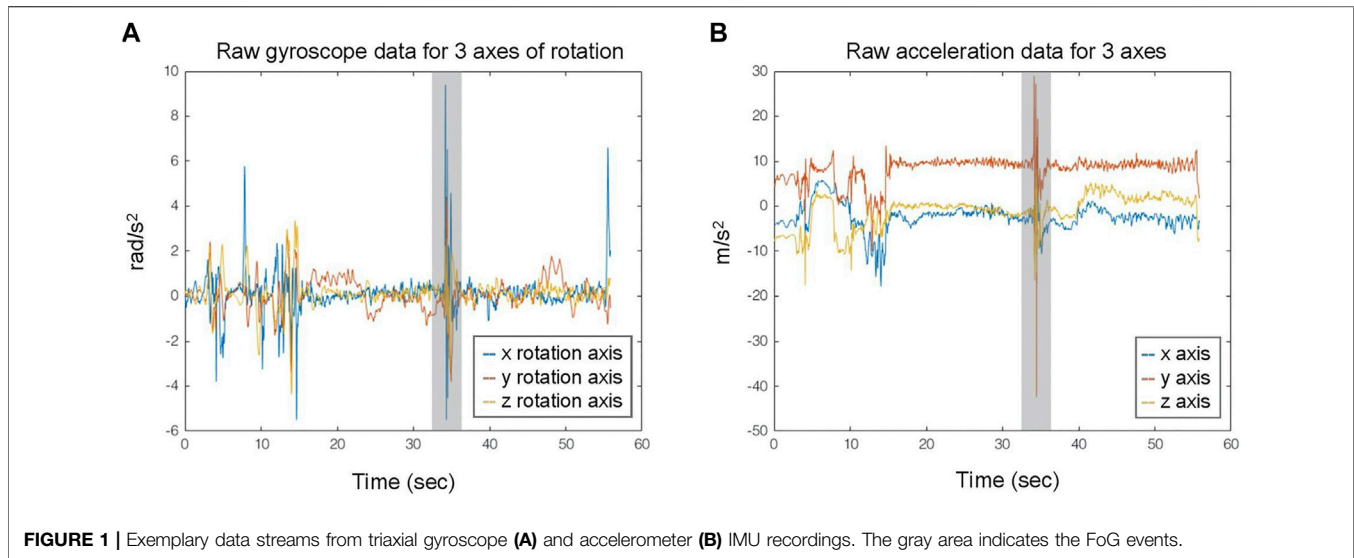
## 2.1 Description of the Database

The CuPiD IMU dataset (Bachlin et al., 2009) contains 3-axis accelerometer and gyroscope data from inertial sensors attached on patient wrists (Mazilu et al., 2013), sampled at 128 Hz. Measurements from 18 patients are included with begin and end timestamps of FoGs, Stops, and Walking-with-Turns.

Subjects in the CuPiD IMU dataset have an age range from 49 to 89 years old (mean: 68.9 years, standard deviation: 10.2 years) and have been diagnosed with PD from 2 to 18 years (mean: 8.8 years, standard deviation: 4.6 years). Each patient performed a walking session with 180° and 360° turns, in wide or narrow trails with obstacles that stimulate FoG events in a controlled environment. During the session, subjects were being asked to perform tasks and walk through crowded hospital rooms. In total, 184 FoG episodes are labeled from clinicians in 11 out of 18 patients, referred as S1–S11. The data of the seven subjects that did not display any episodes during the protocol are not included in the analysis. Time periods in which patients stood and discussed with clinicians or voluntarily stopped are also included in the dataset and are labeled as "Stop." Labeled FoG events in the dataset can occur from 0.11 s up to 98.8 s (mean: 9.12 s, standard deviation: 15.35 s). Most episodes are short, with 50.8% of them lasting less than 3 s and 64.7% less than 5 s. FoG duration range is an important factor for detecting the events, as it plays an important role in defining the interval of the window and the slide-step.

An excerpt from the accelerometer and gyroscope data acquired from the S2 is depicted in **Figure 1**. In the latter, the occurrence of a FoG event (denoted with gray area) is evident near the 35s, exhibiting a distinguishable pattern during the walking activity, simultaneously appearing in both types of IMU data at 35s.

## 2.2 Preprocessing

Data three-dimensional (3D) IMU measurements are extracted for every PD patient from the CuPiD dataset. In particular, for each PD patient, the corresponding $N$-sample accelerometer and gyroscope

**FIGURE 1** | Exemplary data streams from triaxial gyroscope **(A)** and accelerometer **(B)** IMU recordings. The gray area indicates the FoG events.

data (sampled at $f_s$ = 128 Hz), i.e., ($\{\mathbf{a}_x: [a_x[1], a_x[2],\ldots, a_x[N]]^T\}$, $\{\mathbf{a}_y: [a_y[1], a_y[2],\ldots, a_y[N]]^T\}$, $\{\mathbf{a}_z: [a_z[1], a_z[2]\ldots, a_z[N]]^T\}$) and ($\{\mathbf{g}_x: [g_x[1], g_x[2],\ldots, g_x[N]]^T\}$, $\{\mathbf{g}_y: [g_y[1], g_y[2],\ldots, g_y[N]]^T\}$, $\{\mathbf{g}_z: [g_z[1], g_z[2]\ldots, g_z[N]]^T\}$) respectively, were constructed. The data are processed and formatted as a $N \times 6$ matrix for every PD patient, i.e., $S_j = [\mathbf{a}_x,\mathbf{a}_y,\mathbf{a}_z,\mathbf{g}_x,\mathbf{g}_y,\mathbf{g}_z]$, $j = 1,\ldots, 11$ (PD patients).

Further, a three-second ($W_l$=3 s) sliding window with 0.25 s overlap is being selected for processing the $S_j$ data. The duration of the window and the overlap have been chosen as a trade-off between the duration of FoG events and the time resolution needed to detect a differentiation in wrist movement (Mazilu et al., 2015a). Every element of $S_j$ is being normalized with the maximum value per vector across all PD patients and rounded to 4 decimal digits for smaller bit storage.

## 2.3 Segmentation and Training Set Construction

The resulted $S_j$ matrix is being reshaped to $K \times W_l \times 6$, where $K = \frac{N-W_l}{S_l} + 1$ is the number of segments created. These segments were used to construct the training set for the CNN, labeled as "FoG" whether the FoG episode exists inside a window and whether a FoG event starts and ends inside this window. This was adopted so as to capture even short FoG events ($\leq W_l/2$). In the case where a FoG event is not completely integrated inside a window, the segment is being labeled as "FoG" only when the duration of the FoG is $\geq W_l/2$. **Figure 2** illustrates the training set construction process.
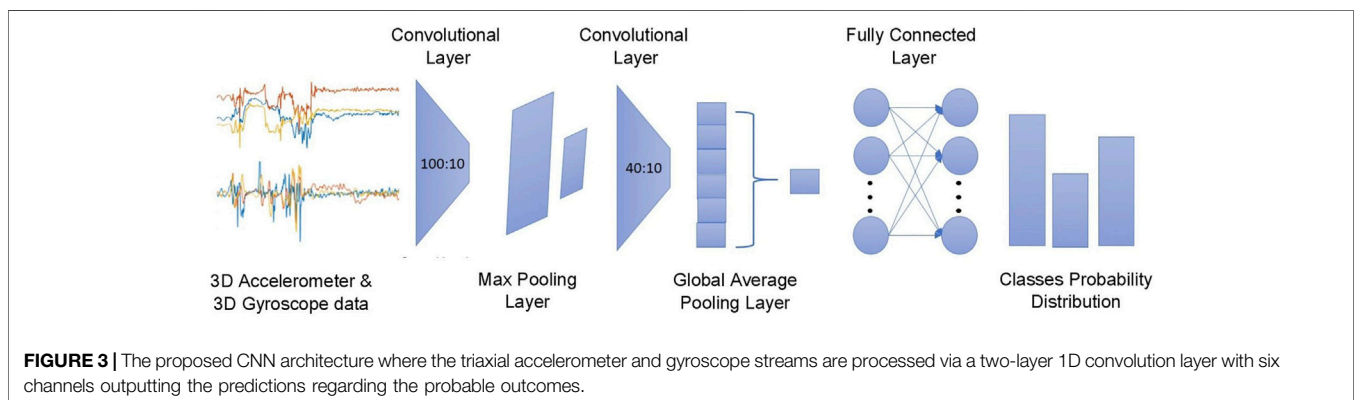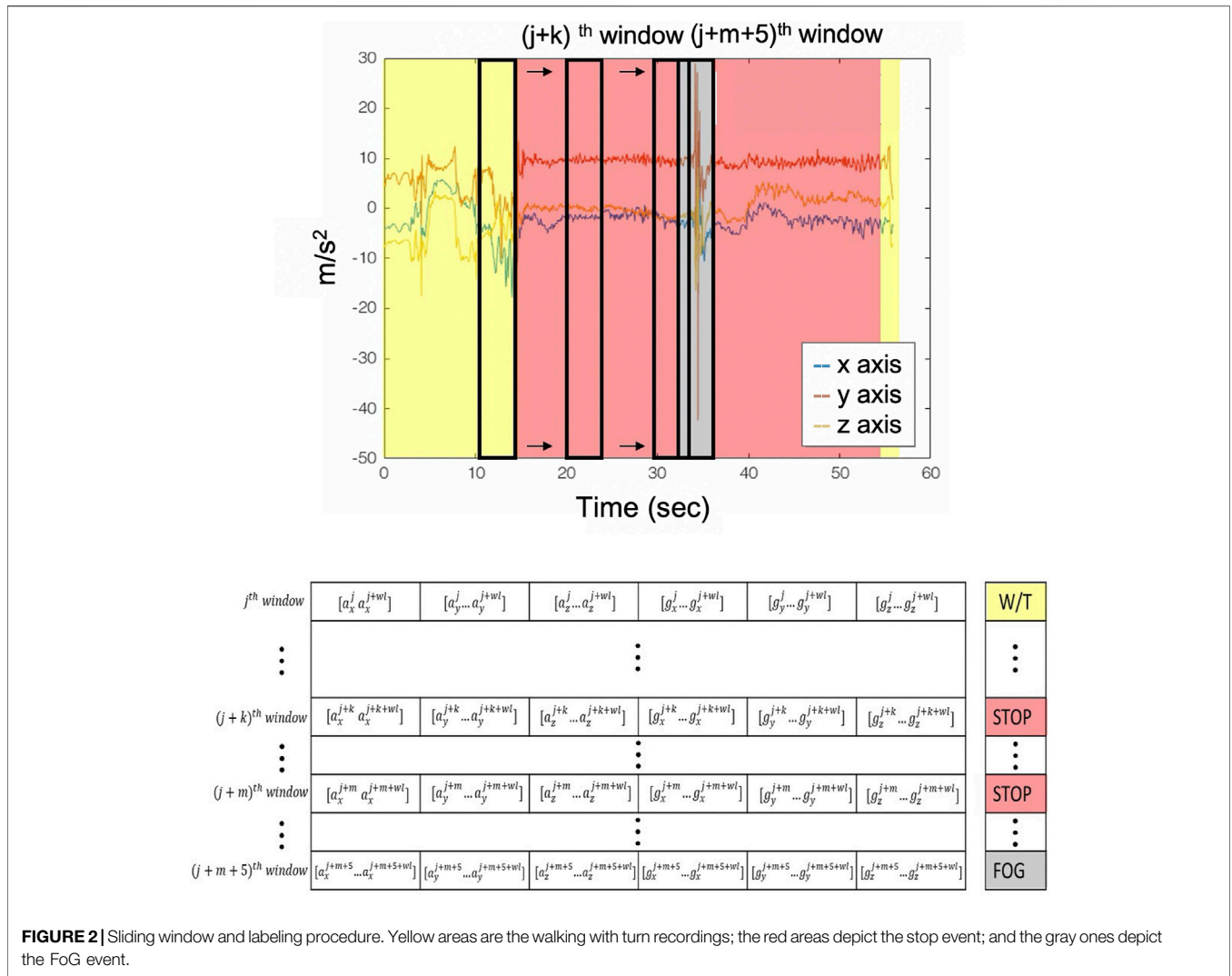
## 2.4 Hyperparameter Tuning and Training of the CNN

Different CNN architectures are assessed by searching the best set of hyperparameters. The hyperparameters that were optimized include the depth of the network$\{2,3,4,5\}$, different kernel sizes $[3–12]$, and number of filters $[10–100]$. The resulting CNN architecture, as depicted in **Figure 3**, consists of a two-layer one-dimensional (1D) convolution that processes the matrix $S_j$ with 100 and 40

filters, respectively, and a kernel size of 10, followed by a fully connected layer with an output of three nodes (i.e., "FoG," "Stop," "Walking-with-turns"). The first convolutional layer is being followed by a max-pooling layer, keeping the maximum value out for every three weights, in order to reduce the complexity and avoid over-fitting. For the second pooling layer, instead of the maximum, the average-pooling is used and the value of two weights is being taken. In this respect, only one weight remains per feature detector and the over-fitting is further prevented (Cook, 2017). Before the dense layer, a 50% dropout is applied to make the model intolerant to small abnormalities of the data (Srivastava et al., 2014). The CNN network is optimized to minimize the categorical cross entropy loss using the ADAM optimizer (Kingma and Ba, 2014). To mitigate overfitting, the model was trained for optimal epochs values. For selecting the optimal value for epochs, the train and validation losses were calculated for each of the CNN models. Loss values were monitored by Early stopping call back function. When there was an increment observed in loss values, training came to halt and the respective value of epoch indicated the optimal selection.

## 2.5 Evaluation Scheme

To evaluate the performance of the proposed model, two approaches are investigated. Considering that the model detects a FoG episode whenever a single window is being classified, sensitivity and specificity of the FoG class are calculated in a 10-CV and LOSO CV manner, so as to test the ability of the network to detect FoG events on unseen data. In the case of 10-CV, the dataset is shuffled and then split into 10 groups. One group acts as the test dataset, while the rest is being used for training. This procedure is repeated 10 times for each group and the average metrics are calculated to describe the results of the evaluation method. The adopted LOSO CV scheme follows the same procedure as the 10-CV with the difference being that the dataset is not randomly split, but every group holds the data of a single patient, with the groups being 11 each for S1–S11, accordingly.

**FIGURE 2 |** Sliding window and labeling procedure. Yellow areas are the walking with turn recordings; the red areas depict the stop event; and the gray ones depict the FoG event.



**FIGURE 3 |** The proposed CNN architecture where the triaxial accelerometer and gyroscope streams are processed via a two-layer 1D convolution layer with six channels outputting the predictions regarding the probable outcomes.

## 3 RESULTS

**Figure 4** aggregates the metrics for the performance assessment of the two evaluation schemes, i.e., 10-CV and LOSO-CV. The

comparisons between the CNN and the traditional classifiers are also presented below in **Table 2**. The specificity and sensitivity were reported to be equal to 88 and 83%, for LOSO-CV, and 90 and 86%, for 10-CV, respectively. Results for patient-specific
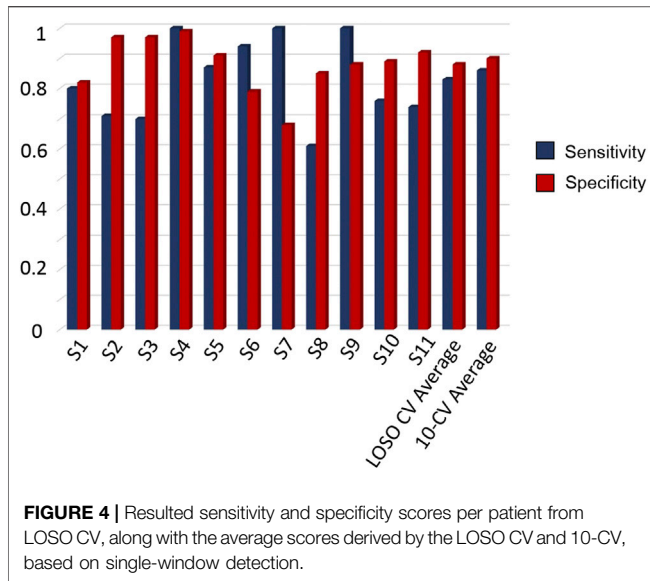
**FIGURE 4 |** Resulted sensitivity and specificity scores per patient from LOSO CV, along with the average scores derived by the LOSO CV and 10-CV, based on single-window detection.

**TABLE 2 |** Single-wrist comparison of sensitivity/specificity pairs for different analysis methodology. F: feature vector consisting of features of Mazilu et al. (2015b).

|  | 10-CV | | LOSO CV | |
|---|---|---|---|---|
|  | **Specificity** | **Sensitivity** | **Specificity** | **Sensitivity** |
| F+Decision Tree | 76% | 71% | 79% | 63% |
| F+XGBoost | 79% | 81% | 83% | 70% |
| CNN (DeepFoG) | 90% | 86% | 88% | 83% |

Deep Learning in the accurate sensor-based detection of FoG episodes. Nonetheless, the use of a waist-placed sensor is likely to impose limitations related to the motion artifacts and inconvenience to the user. It is to be stressed that the comparison of our proposed method with the prior art cannot be direct and absolute, due to the different nature of the required inputs. Concretely, prior works rely on either multiple joint sensors or a single waist-placed sensor.

sensitivity/specificity are presented in **Figure 4**. From the latter, it is observed that some patients' FoGs present different trends in accelerometer and gyroscope data; hence, CNN performs poorly in recognizing them (61% sensitivity, S8 in **Figure 4**).

## 3.1 Performance of the Traditional Classifiers

To validate the assumption that a Deep Learning approach could result in better performance than traditional Machine Learning classifiers in detecting whether a window is FoG or not, this study compares the proposed CNN's performance in training/testing with the performance of the Decision Trees and Extreme Boosting (XGboost) classification algorithms, incorporating the feature set proposed by Mazilu et al. (2015b) but coming from the single-wrist data. A higher performance is noticed from the CNN architecture (*see* sensitivity/specificity pairs in **Table 2**), hence, it is more suitable to be utilized for single-wrist detection.

## 3.2 Comparison of CNN With State-of-The-Art

Comparison between the DeepFoG performance and previously published approaches (with and without sensors mounted to the wrists of the patients) is presented in **Table 1**. Direct comparison is more relevant to the methods using sensors in the upper limb (Cole et al., 2011; Tripoliti et al., 2013; Mazilu et al., 2015b). It was demonstrated that DeepFoG has comparable classification performance with the related works in the literature which, however, use multiple sensors. The work by Camps et al. is of importance. The authors were the first to introduce a Deep Learning method for detecting FoG episodes in PD patients (Camps et al., 2018). In their approach, they used data collected by a waist-placed IMU (21 PD patients). Their proposed methodology achieved sensitivity/specificity equal to 92.6/88.7%. This study set a promising example for the use of

## 4 DISCUSSION

Digital Health is an emerging field that could assist PD management via the realization of accessible tools towards detection and monitoring of the disease. Moreover, FoG episodes are strongly associated with the increased risk of fall and frailty of PD patients. Ultimately, the output of the trained DeepFoG, proposed in the current study, can then inform a timely application of RAS, combined with a hand vibration stimulus, to facilitate the PD patient to overcome the FoG event. The use of a smartwatch and the distinction between FoG events and the different natural movements of the user enables the transferability of the solution to a real-life environment. The latter reinforces the added value of the results towards capturing PD FoG episodes and provides objective information to the doctor regarding the frequency of the FoG event occurrence, towards objective monitoring of the PD patient. Moreover, the unobtrusive character of DeepFoG data acquisition contributes to the achievement of long-term adherence to the proposed solution, and enables the regulation of the medication towards reducing PD related risks.

DeepFoG may not outperform the related State-of-the-Art methods in both sensitivity and specificity metrics (**Table 1**). However, high performance is achieved by using only a single device (namely, smartwatch), when the others need at least two to arrive at a similar sensitivity/specificity. FoG events can be considered as rare events in the sensor streams during walking, so DeepFoG's high specificity enables the ecological validity of the solutions. Moreover, the latency of DeepFoG to infer the FoG event is near real-time (considering 0.25 s time-window for each decision outcome); hence, even this retrospective study allows envisioning the timely automatic-RAS/hand vibration interventions. The DeepFoG can be integrated in a fused scheme with more PD risk-related variables, like heart rate variability features, age, or history of falls and FoGs, via a fuzzy-logic based model (Iakovakis et al.,

2016), that could formulate an assisting feedback mechanism tailored to the everyday living activities of high-risk PD patients.

From a Deep Learning perspective, the use of the LOSO evaluation scheme informs us regarding the generalizability of the CNNs to detect FoG events for new subjects. It can also be observed that the 10-CV results were improved by 3 and 2% in specificity and sensitivity, when compared to the LOSO evaluation. Moreover, from the patient specific sensitivity/ specificity pairs depicted in **Figure 4**, it can be observed that personalized cuing and FoG episodes could be detected via the use of personalized training for tuning the networks parameters to capture the patient-specific characteristics. The proposed version of DeepFoG was designed to classify three classes, "Walking with turns," "Stop,", and "FoG" achieving it with high $F$-score 0.81; instead of distinguishing just "FoG" and "Non FoG." This makes DeepFoG insusceptible to false alarms FoG events due to intentional stopping, enhancing in this way its user friendliness and acceptability.

Despite the relatively high specificity and sensitivity estimated metrics, the dataset used to train and evaluate the CNN is quite limited (11 patients with FoG) and the FoG episodes are few (184). Hence, generalization of the outcomes should be performed with caution. Nevertheless, the recordings were longitudinal during the day and the detection results can be considered relevant and significant to encourage the continuation of the research towards evaluation of FoG detection-intervention using data from wearable devices. Importantly, we recognize that the optimal way to test a model's generalization ability is by externally testing the model using an entirely new dataset. Given that such data are not currently available, we chose to perform a preliminary validation of the proposed framework before proceeding with a larger cohort. Provided that the preliminary evaluation gives promising results, we envision to continue with a larger population in our future research. Furthermore, lower prediction outcomes in some patients may raise questions about the various FoG patterns that exist in the examined cases and, thus, different gait patterns and their change along the PD severity need further investigation. Pilot tests on PD patients should be initialized in order to evaluate the quality and efficiency of the proposed system in home-based daily circumstances. Research effort towards this direction is already in progress.

## 5 CONCLUSION

The current study introduces a single-wrist-based FoG detection scheme that incorporates IMU data from a commercial smartwatch and CNN to facilitate a home-based falls prevention and FoG management. Instantaneous risk estimation enables the reduction of PD-related risks and can sustain the patient's QoL. The Deep Learning-based method of single-arm might achieve similar accuracy as previously published methods, but with utilization of fewer sensors. The main advantage offered by our proposed methodology is its simplification and convenience attributed to the use of a single smartwatch rather than its improved accuracy. Nevertheless, our approach could be more preferable and easier accepted by the patients compared to the State-of-the-Art for feature extraction and classification methodology. Further evaluation of the proposed methodology on a large PD cohort in high risk remains to be conducted. Upon successful clinical validation, this work may provide an objective method to inform the doctors regarding the frequency of FoG events and it could test the effectiveness of the timely RAS/hand vibration interventions.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## AUTHOR CONTRIBUTIONS

TB, DI, and LH conceived the pipeline; TB developed the algorithms, analysed the data, generated figures and tables. TB, DI, and LH contributed to the manuscript. All authors discussed the results.

## FUNDING

## REFERENCES

Allen, N. E., Canning, C. G., Sherrington, C., Lord, S. R., Latt, M. D., Close, J. C., et al. (2010). The Effects of an Exercise Program on Fall Risk Factors in People with Parkinson's Disease: a Randomized Controlled Trial. *Movement Disord.* 25 (9), 1217–1225. doi:10.1002/mds.23082

Bachlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J. M., Giladi, N., et al. (2009). Wearable Assistant for Parkinson's Disease Patients with the Freezing of Gait Symptom. *IEEE Transactions on Information Technology in Biomedicine* 14 (2), 436–446. doi:10.1109/TITB.2009.2036165

Bloem, B. R., Hausdorff, J. M., Visser, J. E., and Giladi, N. (2004). Falls and Freezing of Gait in Parkinson's Disease: a Review of Two Interconnected, Episodic Phenomena. *Movement Disord. official J. Movement Disord. Soc.* 19 (8), 871–884. doi:10.1002/mds.20115

Camps, J., Sama, A., Martin, M., Rodriguez-Martin, D., Perez-Lopez, C., Arostegui, J. M. M., et al. (2018). Deep Learning for Freezing of Gait Detection in Parkinson's Disease Patients in Their Homes Using a Waist-Worn Inertial Measurement Unit. *Knowledge-Based Syst.* 139, 119–131. doi:10.1016/j.knosys.2017.10.017

Cole, B. T., Roy, S. H., and Nawab, S. H. (2011). Detecting Freezing-Of-Gait during Unscripted and Unconstrained Activity. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2011, 5649–5652. doi:10.1109/IEMBS.2011.6091367

Cook, A. (2017). Global Average Pooling Layers for Object Localization. Available at: https://alexisbcook. github. io/2017/global-average-pooling-layers-forobject-localization

Coria, F., and Cozar-Santiago, M. d. P. (2008). Rasagiline Improves Freezing in a Patient with Primary Progressive Freezing Gait. *Movement Disord. official J. Movement Disord. Soc.* 23 (3), 449–451. doi:10.1002/mds.21868

Dalla Bella, S., Benoit, C.-E., Farrugia, N., Keller, P. E., Obrig, H., Mainka, S., et al. (2017). Gait Improvement via Rhythmic Stimulation in Parkinson's Disease Is Linked to Rhythmic Skills. *Scientific Rep.* 7, 42005. doi:10.1038/srep42005

Dorsey, E., Constantinescu, R., Thompson, J., Biglan, K., Holloway, R., Kieburtz, K., et al. (2007). Projected Number of People with Parkinson Disease in the Most Populous Nations, 2005 through 2030. *Neurology* 68 (5), 384–386. doi:10.1212/01.wnl.0000247740.47667.03

Fasano, A., Laganiere, S. E., Lam, S., and Fox, M. D. (2017). Lesions Causing Freezing of Gait Localize to a Cerebellar Functional Network. *Ann. Neurol.* 81 (1), 129–141. doi:10.1002/ana.24845

Fietzek, U. M., Zwosta, J., Schroeteler, F. E., Ziegler, K., and Ceballos-Baumann, A. O. (2013). Levodopa Changes the Severity of Freezing in Parkinson's Disease. *Parkinsonism Relat. Disord.* 19 (10), 894–896. doi:10.1016/j.parkreldis.2013.04.004

Giladi, N., Treves, T., Simon, E., Shabtai, H., Orlov, Y., Kandinov, B., et al. (2001). Freezing of Gait in Patients with Advanced Parkinson's Disease. *J. Neural Transm.* 108 (1), 53–61. doi:10.1007/s007020170096

Hall, J. M., Shine, J. M., O'Callaghan, C., Walton, C. C., Gilat, M., Naismith, S. L., et al. (2015). Freezing of Gait and its Associations in the Early and Advanced Clinical Motor Stages of Parkinson's Disease: a Cross-Sectional Study. *J. Parkinson's Dis.* 5 (4), 881–891. doi:10.3233/JPD-150581

Hausdorff, J. M., Lowenthal, J., Herman, T., Gruendlinger, L., Peretz, C., and Giladi, N. (2007). Rhythmic Auditory Stimulation Modulates Gait Variability in Parkinson's Disease. *Eur. J. Neurosci.* 26 (8), 2369–2375. doi:10.1111/j.1460-9568.2007.05810.x

Hauser, R. A. (2009). Levodopa: Past, Present, and Future. *Eur. Neurol.* 62 (1), 1–8. doi:10.1159/000215875

Heremans, E., Nieuwboer, A., and Vercruysse, S. (2013). Freezing of Gait in Parkinson's Disease: where Are We Now? *Curr. Neurol. Neurosci. Rep.* 13 (6), 350. doi:10.1007/s11910-013-0350-7

Huang, C., Chu, H., Zhang, Y., and Wang, X. (2018). Deep Brain Stimulation to Alleviate Freezing of Gait and Cognitive Dysfunction in Parkinson's Disease: Update on Current Research and Future Perspectives. *Front. Neurosci.* 12, 29. doi:10.3389/fnins.2018.00029

Iakovakis, D. E., Papadopoulou, F. A., and Hadjileontiadis, L. J. (2016). Fuzzy Logic-Based Risk of Fall Estimation Using Smartwatch Data as a Means to Form an Assistive Feedback Mechanism in Everyday Living Activities. *Healthc. Technol. Lett.* 3 (4), 263–268. doi:10.1049/htl.2016.0064

Jovanov, E., Wang, E., Verhagen, L., Fredrickson, M., and Fratangelo, R. (2009). Defog--A Real Time System for Detection and Unfreezing of Gait of Parkinson's Patients. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2009, 5151–5154. doi:10.1109/IEMBS.2009.5334257

Kingma, D. P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980

Lim, I., van Wegen, E., de Goede, C., Deutekom, M., Nieuwboer, A., Willems, A., et al. (2005). Effects of External Rhythmical Cueing on Gait in Patients with Parkinson's Disease: a Systematic Review. *Clin. Rehabil.* 19, 695–713. doi:10.1191/0269215505cr906oa

Mazilu, S., Blanke, U., Dorfman, M., Gazit, E., Mirelman, A. M., Hausdorff, J., et al. (2015a). A Wearable Assistant for Gait Training for Parkinson's Disease with Freezing of Gait in Out-Of-The-Lab Environments. *ACM Trans. Interactive Intell. Syst. (Tiis)* 5, 1–31. 10.1145/2701431

Mazilu, S., Blanke, U., Roggen, D., Tröster, G., Gazit, E., and Hausdorff, J. M. (2013). Engineers Meet Clinicians: Augmenting Parkinson's Disease Patients to Gather Information for Gait Rehabilitation. in AH '13: Proceedings of the 4th Augmented Human International Conference, 124–127. doi:10.1145/2459236.2459257

Mazilu, S., Blanke, U., and Tröster, G. (2015b). Gait, Wrist, and Sensors: Detecting Freezing of Gait in Parkinson's Disease from Wrist Movement. in 2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops) (IEEE), 579–584.

Mazilu, S., Hardegger, M., Zhu, Z., Roggen, D., Tröster, G., Plotnik, M., et al. (2012). 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 (IEEE), 123–130.

Niazmand, K., Tonn, K., Zhao, Y., Fietzek, U., Schroeteler, F., Ziegler, K., et al. (2011). Freezing of Gait Detection in Parkinson's Disease Using Accelerometer Based Smart Clothes. in 2011 IEEE Biomedical Circuits and Systems Conference (BioCAS) (IEEE), 201–204.

Nutt, J. G., Bloem, B. R., Giladi, N., Hallett, M., Horak, F. B., and Nieuwboer, A. (2011). Freezing of Gait: Moving Forward on a Mysterious Clinical Phenomenon. *Lancet Neurol.* 10 (8), 734–744. doi:10.1016/S1474-4422(11)70143-0

Palmerini, L., Rocchi, L., Mazilu, S., Gazit, E., Hausdorff, J. M., and Chiari, L. (2017). Identification of Characteristic Motor Patterns Preceding Freezing of Gait in Parkinson's Disease Using Wearable Sensors. *Front. Neurol.* 8, 394. doi:10.3389/fneur.2017.00394

Pozzi, N. G., Canessa, A., Palmisano, C., Brumberg, J., Steigerwald, F., Reich, M. M., et al. (2019). Freezing of Gait in Parkinson's Disease Reflects a Sudden Derangement of Locomotor Network Dynamics. *Brain* 142 (7), 2037–2050. 10.1093/brain/awz141

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks From Overfitting. *J. machine Learn. Res.* 15, 1929–1958.

Thaut, M. H., McIntosh, G. C., Rice, R. R., Miller, R. A., Rathbun, J., and Brault, J. (1996). Rhythmic Auditory Stimulation in Gait Training for Parkinson's Disease Patients. *Movement Disord. official J. Movement Disord. Soc.* 11 (2), 193–200. doi:10.1002/mds.870110213

Tripoliti, E. E., Tzallas, A. T., Tsipouras, M. G., Rigas, G., Bougia, P., Leontiou, M., et al. (2013). Automatic Detection of Freezing of Gait Events in Patients with Parkinson's Disease. *Computer Methods Programs Biomed.* 110, 12–26. doi:10.1016/j.cmpb.2012.10.016

Vercruysse, S., Spildooren, J., Heremans, E., Vandenbossche, J., Levin, O., Wenderoth, N., et al. (2012). Freezing in Parkinson's Disease: a Spatiotemporal Motor Disorder beyond Gait. *Movement Disord.* 27 (2), 254–263. doi:10.1002/mds.24015

Walton, C. C., Mowszowski, L., Gilat, M., Hall, J. M., O'Callaghan, C., Muller, A. J., et al. (2018). Cognitive Training for Freezing of Gait in Parkinson's Disease: a Randomized Controlled Trial. *NPJ Parkinsons Dis.* 4, 15. doi:10.1038/s41531-018-0052-6

Wang, J., Chen, Y., Hao, S., Peng, X., and Hu, L. (2019). Deep Learning for Sensor-Based Activity Recognition: A Survey. *Pattern Recognition Lett.* 119, 3–11.

Witt, I., Ganjavi, H., and MacDonald, P. (2019). Relationship between freezing of gait and anxiety in parkinson's disease patients: A systemic literature review. *Parkinsons Dis.* 2019, 6836082.

Zhang, L.-L., Canning, S. D., and Wang, X.-P. (2016). Freezing of Gait in Parkinsonism and its Potential Drug Treatment. *Curr. neuropharmacology* 14 (4), 302–306. doi:10.2174/1570159x14666151201190040

Ziv, I., Avraham, M., Dabby, R., Zoldan, J., Djaldetti, R., and Melamed, E. (1999). Early-occurrence of Manual Motor Blocks in Parkinson's Disease: a Quantitative Assessment. *Acta Neurol. Scand.* 99 (2), 106–111. doi:10.1111/j.1600-0404.1999.tb00666.x