



Provided by the author(s) and NUI Galway in accordance with publisher policies. Please cite the published version when available.

Title	Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer
Author(s)	Rodríguez-Martín, Daniel; Samà, Albert; Pérez-López, Carlos; Català, Andreu; Moreno Arostegui, Joan M.; Cabestany, Joan; Bayés, Àngels; Alcaine, Sheila; Mestre, Berta; Prats, Anna; Cruz Crespo, M.; Counihan, Timothy J.; Browne, Patrick; Quinlan, Leo R.; Ó Laighin, Gearóid; Sweeney, Dean; Lewy, Hadas; Azuri, Joseph; Vainstein, Gabriel; Annicchiarico, Roberta; Costa, Alberto; Rodríguez-Molinero, Alejandro
Publication Date	2017-02-15
Publication Information	Rodriguez-Martin, D;Sama, A;Perez-Lopez, C;Catala, A;Arostegui, JMM;Cabestany, J;Bayes, A;Alcaine, S;Mestre, B;Prats, A;Crespo, MC;Counihan, TJ;Browne, P;Quinlan, LR;OLaighin, G;Sweeney, D;Lewy, H;Azuri, J;Vainstein, G;Annicchiarico, R;Costa, A;Rodriguez-Molinero, A (2017) 'Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer'. PLOS ONE, 12(2), e0171764. doi: 10.1371/journal.pone.0171764
Publisher	Public Library of Science
Link to publisher's version	<a href="http://dx.doi.org/10.1371/journal.pone.0171764">http://dx.doi.org/10.1371/journal.pone.0171764</a>
Item record	<a href="http://hdl.handle.net/10379/6424">http://hdl.handle.net/10379/6424</a>
DOI	<a href="http://dx.doi.org/10.1371/journal.pone.0171764">http://dx.doi.org/10.1371/journal.pone.0171764</a>

Downloaded 2020-10-17T01:32:39Z

Some rights reserved. For more information, please see the item record link above.



RESEARCH ARTICLE

# Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer

Daniel Rodríguez-Martín<sup>1</sup>\*, Albert Samà<sup>1,2</sup>, Carlos Pérez-López<sup>1,2</sup>, Andreu Català<sup>1,2</sup>, Joan M. Moreno Arostegui<sup>1,2</sup>, Joan Cabestany<sup>1,2</sup>, Àngels Bayés<sup>3</sup>, Sheila Alcaine<sup>3</sup>, Berta Mestre<sup>3</sup>, Anna Prats<sup>3</sup>, M. Cruz Crespo<sup>3</sup>, Timothy J. Counihan<sup>4</sup>, Patrick Browne<sup>4</sup>, Leo R. Quinlan<sup>5</sup>, Gearóid ÓLaighin<sup>5</sup>, Dean Sweeney<sup>5</sup>, Hadas Lewy<sup>6</sup>, Joseph Azuri<sup>6,7</sup>, Gabriel Vainstein<sup>6</sup>, Roberta Annicchiarico<sup>8</sup>, Alberto Costa<sup>8</sup>, Alejandro Rodríguez-Molinero<sup>2,5</sup>



**1** Universitat Politècnica de Catalunya – BarcelonaTech (UPC), Technical Research Centre for Dependency Care and Autonomous Living (CETPD), Vilanova i la Geltrú, Spain, **2** Sense4Care, Barcelona, Spain, **3** Unidad de Parkinson y trastornos del movimiento (UParkinson), Barcelona, Spain, **4** School of Medicine, National University of Ireland Galway (NUIG), Galway, Ireland, **5** Electrical & Electronic Engineering Department, National University of Ireland Galway (NUIG), Galway, Ireland, **6** Maccabi Healthcare Services, Tel Aviv, Israel, **7** Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel, **8** IRCCS Fondazione Santa Lucia, Rome, Italy

**OPEN ACCESS**

**Citation:** Rodríguez-Martín D, Samà A, Pérez-López C, Català A, Moreno Arostegui JM, Cabestany J, et al. (2017) Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer. PLoS ONE 12(2): e0171764. doi:10.1371/journal.pone.0171764

**Editor:** Sergio Gómez, Universitat Rovira i Virgili, SPAIN

**Received:** July 22, 2016

**Accepted:** January 25, 2017

**Published:** February 15, 2017

**Copyright:** © 2017 Rodríguez-Martín et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Due to ethical restrictions imposed by Spanish Agency for Drugs and Medical Devices (Spain), Clinical Research Ethics Committee-Galway University Hospitals (Ireland), Health Department Maccabi Healthcare Services (Israel) and Ethics Committee Fondazione Santa Lucia (Italy), related to approved consent procedure and protecting privacy, all relevant data are available upon request (Joan Cabestany, email: [joan.cabestany@upc.edu](mailto:joan.cabestany@upc.edu) and Daniel Rodríguez-Martín, email: [daniel.rodriguez-martin@upc.edu](mailto:daniel.rodriguez-martin@upc.edu)).

\* These authors contributed equally to this work.

✉ Current address: Clinical Research Unit, Consorci Sanitari del Garraf, Vilanova i la Geltrú, Spain

\* [daniel.rodriguez-martin@upc.edu](mailto:daniel.rodriguez-martin@upc.edu)

## Abstract

Among Parkinson’s disease (PD) symptoms, freezing of gait (FoG) is one of the most debilitating. To assess FoG, current clinical practice mostly employs repeated evaluations over weeks and months based on questionnaires, which may not accurately map the severity of this symptom. The use of a non-invasive system to monitor the activities of daily living (ADL) and the PD symptoms experienced by patients throughout the day could provide a more accurate and objective evaluation of FoG in order to better understand the evolution of the disease and allow for a more informed decision-making process in making adjustments to the patient’s treatment plan. This paper presents a new algorithm to detect FoG with a machine learning approach based on Support Vector Machines (SVM) and a single tri-axial accelerometer worn at the waist. The method is evaluated through the acceleration signals in an outpatient setting gathered from 21 PD patients at their home and evaluated under two different conditions: first, a generic model is tested by using a leave-one-out approach and, second, a personalised model that also uses part of the dataset from each patient. Results show a significant improvement in the accuracy of the personalised model compared to the generic model, showing enhancement in the specificity and sensitivity geometric mean (GM) of 7.2%. Furthermore, the SVM approach adopted has been compared to the most comprehensive FoG detection method currently in use (referred to as MBFA in this paper). Results of our novel generic method provide an enhancement of 11.2% in the GM compared to the MBFA generic model and, in the case of the personalised model, a 10% of improvement with respect to the MBFA personalised model. Thus, our results show that a machine

**Funding:** This project has been performed within the framework of the "Freezing in Parkinson's Disease: Improving Quality of Life with an Automatic Control System" (MASPARK) project which is funded by La Fundació La Marató de TV3 20140431 [[http://www.tv3.cat/marato/es/projectes\\_financats\\_2013](http://www.tv3.cat/marato/es/projectes_financats_2013)]. This work also forms part of the framework of the FP7 project "Personal Health Device for the Remote and Autonomous Management of Parkinson's Disease" (REMPARK) ICT-287677, which is funded by the European Community [<http://www.rempark.eu/>].

**Competing interests:** Albert Samà, Carlos Pérez-López, Andreu Català, Juan Manuel Moreno, Joan Cabestany and Alejandro Rodríguez-Molinero are shareholders of Sense4Care, which is a spin-off company. Sense4Care may commercialize the results of this research device in a near future. These authors declare that the possible commercialization of the product is a research outcome, not being the design, the analysis, the interpretation of the results or the conclusions affected by commercial interests. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

learning approach can be used to monitor FoG during the daily life of PD patients and, furthermore, personalised models for FoG detection can be used to improve monitoring accuracy.

## Introduction

Parkinson's disease (PD) is a neurodegenerative disease that principally affects the motor system. According to the Global Declaration for Parkinson's Disease, PD affects up to 6.3 million people worldwide [1]. Among the many PD symptoms, Freezing of Gait (FoG) is one of the most incapacitating and is usually present in the more advanced phase of the disease [2,3]. FoG is commonly described by PD patients as if their feet were "glued to the floor" resulting in loss of postural balance frequently causing falls [4,5]. In addition to the motor complications arising from FoG, it can also lead to non-motor complications including social isolation, depression, and anxiety. [6,7].

The precise tracking of the occurrence of FoG is challenging for clinicians. The acquisition of such information can greatly help in optimising therapies including pharmacotherapy, and physiotherapy, that are known to be beneficial for reducing FoG episodes [4]. Furthermore, by acquiring accurate information on FoG frequency and severity in conjunction with the full repertoire of other motor and non-motor symptomatology, this would provide to the neurologists a more accurate overall view of the disease status and evolution over time. Additionally, the online detection of FoG opens up the possibility of incorporating actuation methodologies such as rhythmic auditory cues that might shorten or prevent FoG episodes [8,9]. These rhythmic auditory cues are hereafter referred to as *cueing*.

In clinical practice, there are two methods employed to evaluate the presence of FoG. The first is based on specific gait tests, for example the Timed Up & Go test performed through a narrow space, using specific circuits with several turns [10–12] or by assessing patients while performing rapid and short steps [13]. The second method is based on answers given by patients and caregivers about the frequency and severity of FoG episodes collected in specific FoG questionnaires [14–16]. These methods have major drawbacks due to both the clinical setting not reflecting real life settings as well as the timing of the evaluation not replicating the recognised variability in FOG frequency. FoG episodes often occur in the patient's own home and their frequency is reduced significantly in controlled environments such as the laboratory setting [17]. Self-assessment of FoG, has shown to be unreliable since patients and caregivers may not perceive the episodes and, furthermore, because PD patients often experience memory loss, inattention or dementia, leading to imprecise recall [18].

Therefore, many attempts have been made to obtain an objective and precise method for monitoring FoG in patients during their normal activities of daily living (ADL). In this context, many different types of sensors have been employed. The most widespread method to objectively detect FoG is based on the analysis of data emanating from wearable motion sensors, as they are unobtrusive, portable with low power consumption, providing currently the most suitable way to evaluate FoG episodes in patients' own homes.

Most methods to detect FoG based on wearable sensors have been developed and tested under controlled conditions [19–27], which is mainly due to the difficulty of obtaining a valid gold standard outside the laboratory setting. Therefore, trained algorithms are likely to generate several false positives when they are tested in real life since algorithms have not been trained for specific situations that could happen in patients daily living activities. As far as the

authors are aware, only Ahlrichs et al. have generated an algorithm to detect FoG in PD patients' own homes [28]; however, part of the evaluation was not performed with PD patients with FoG and the latency time of the algorithm was every minute, which is too slow to be effective for cueing purposes. Until now, objective FoG assessment during ADLs in the home has not been extensively examined, thus, it might not be reliable. The usability and acceptability of any new wearable device is critical to its success. Many studies have employed several sensors at specific locations (Table 1) providing promising results in some cases. However, their usage may be counterproductive when monitoring FoG at home, as they require patients to wear many sensors, reducing the usability of the monitoring system and its acceptability. Thus, the trade-off between accuracy and usability is a key factor in designing such systems. The waist is one of the most suitable placements for detecting FoG [28–31]. Furthermore, according to Mathie et al., locating a wearable system at a side of the waist is the most comfortable place, as concluded in a questionnaire performed by elderly people [32]. Additionally, in humans the waist is close to the centre of gravity better representing body movement [33] and also allowing the sensor to monitor other PD symptoms such as bradykinesia or dyskinesia [34,35].

This paper presents a study performed within the 'Freezing in Parkinson's Disease: Improving Quality of Life with an Automatic Control System' (MASPARK) project [36] for detecting FoG by locating a single wearable sensor at the waist. More specifically, this paper presents a new machine learning based algorithm to detect FoG episodes validated by clinical experts, in PD patients' own home [37]. Furthermore, and with the aim of evaluating the algorithm in real life, the proposed algorithm is also designed with the purpose of being implemented in a microcontroller, thus, some issues in the computational resources used by the method such as windowing and the classification cost have also been taken into account. The approach presented here is compared to the most extended algorithm used to detect FoG (hereafter referred to as the *Moore-Bächlin FoG Algorithm*, MBFA) [21]. The validation methodology employed for both approaches consists of a Leave-One-Patient-Out. Furthermore, this leave one out analysis is compared to a user-dependent algorithm, in which an individualised model classifier is evaluated. Results on 21 PD patients show that the generic proposed approach outperforms the current most employed FoG detection method by an average 11%, with a sensitivity of 74.7% (proportion of actual positives that are correctly identified as positives) and specificity of 79% (proportion of actual negatives correctly identified as negatives) compared to a sensitivity of 81.6% and specificity of 52.6%. In addition, personalisation of the detection model achieves an improvement in the symptom detection of around 10% on average (80.1 on sensitivity, 88.1% on specificity against an 89.1% on sensitivity and 61.5% on specificity). Thus, a machine learning approach can be used to monitor FoG during the daily life of PD patients and, furthermore, personalised models for FoG detection can be used to improve the monitoring accuracy.

The paper is organised as follows: first a background with the most relevant FoG detection systems and algorithms is introduced, then the proposed methodologies, the experiments performed and the evaluation of the algorithms are presented, and finally, the results, discussion and conclusions are provided.

## Related work

FoG has been analysed with a wide variety of systems and sensors. Some of these reports have employed systems to capture FoG episodes that are not transferable to normal daily life of patients since they can only be used in a laboratory setting. Some examples are pressure platforms [38–41], which are non-portable systems, electroencephalogram (EEG) [42], electromyography (EMG) [25] or skin conductance [43], which involves the placement of electrodes on

**Table 1. Summary of literature regarding patients and sensor systems.**

Work	Year	# Patients			Sensors	# Sensors	Location
		PD-Freezers*	PD Non-Freezers	Control			
Nieuwboer et al. [39]	2001	17(14)	-	-	Camera, Force Plates, EMG	6, 1 and 6 respectively	EMG at lower limbs, the others are not wearable
Han et al. [19]	2003	2(2)	-	-	Accelerometers	2	2 x ankle
Hausdorff et al. [38]	2003	11(11)	-	-	Insole pressure	1	Not wearable
Moore et al. [20]	2008	11(7)	-	-	Accelerometers	1	Left ankle
Daphnet project [56]	2009–2011	10 (8)	-	-	Accelerometers	1	Knee
Zabaleta et al. [24]	2008	2(2)	-	-	Accelerometers	6	2 x ankle, 2 x knee, 2 x thigh
Jovanov et al. [22]	2009	1(1)	-	4	Accelerometers	1	Ankle
Maidan et al. [57]	2010	10(10)	10	15	ECG, Insoles and acc.	2 each	ECG leads at chest, insole and acc. at shoe.
Knobl et al. [40]	2010	15(15)	16	16	Pressure platform	1	Not wearable
Delval et al. [23]	2010	10(5)	10	10	Goniometer, Video system	2	2 x Knee
Cole et al. [25]	2011	10(4)	-	2	EMG and Accelerometers	1 and 3 respectively	EMG at shin. Acc. at shin, thigh and forearm
Niazmand et al. [26]	2011	6(6)	-	-	Accelerometers	5	2 x Shank, 2 x Thigh and 1 x lower abdomen
Stamatakis et al. [51]	2011	1(1)	-	1	Accelerometers	4	2 x Hallux and 2 x heel
Almeida et al. [40]	2012	10(10)	10	10	Pressure platform	1	Not wearable
Zhao et al. [46]	2012	8(6)	-	-	Accelerometers	5	2 x shank, 2 x thigh and 1 x lower abdomen
Handojoseno et al. [42]	2012	26(10)	-	-	Encephalogram	4 electrodes	Head
Mancini et al. [47]	2012	21(21)	27	21	Accelerometers	3	1 x Posterior trunk, 2 x shank
Mazilu et al. [48]	2012	10(8)	-	-	Accelerometers	3	Shank, thigh and lower back
Takac et al. [58]	2013	1(1)	-	12	Depth Sensor Camera, Accelerometers	1 each	Acc at waist
Moore et al. [27]	2013	25 (20)	-	-	Accelerometers	7	Lower back, 2xThigh, 2xShank and 2xHeel
Tripoliti et al. [52]	2013	5(5)	6	5	Accelerometers and gyroscopes	6 and 2 respectively	Acc: (2 x shank, 2 x wrist, 1 x chest, 1 x abdomen) Gyro: abdomen and chest
Mazilu et al. [49]	2013	10(8)	-	-	Accelerometers	3	Shank, thigh and lower back
Rodríguez et al. [50]	2014	10(10)	10	-	Accelerometers	1	Waist
Coste et al. [53]	2014	4(4)	-	-	Accelerometer	1	Ankle
Weiss et al. [30]	2015	28(28)	44	-	Accelerometer	1	Lower Back
Tay et al. [54]	2015	8(5)	-	-	Gyroscope	2	2 x Ankle
Zach et al. [31]	2015	23(23)	-	-	Accelerometers	1	Waist
Mazilu et al. [43]	2015	18(11)	-	-	ECG and Skin Conductance	1 each	ECG at chest, SC in 2 fingers
Ahlich et al. [28]	2015	8(8)	12	-	Accelerometers	1	Waist

(Continued)

Table 1. (Continued)

Work	Year	# Patients			Sensors	# Sensors	Location
		PD-Freezers*	PD Non-Freezers	Control			
Maidan et al. [59]	2015	11 (11)	-	11	near Infrared Spectroscopy	1	Head

\* In brackets the total of PD patients who suffered FoG episodes during the tests

doi:10.1371/journal.pone.0171764.t001

the skin in addition to a datalogger system to capture these data. Other obtrusive system examples are knee-joint goniometers [23] or cameras and video systems, which also have a low acceptability by users in a non-laboratory setting [23,39,44]. Thus, given that PD monitoring should be ambulatory and should last several hours in order to provide useful clinical information [34,45], most works have employed non-invasive systems such as wearable systems based in microelectromechanical systems (MEMS).

In 2003, Han et al. used MEMS based inertial systems, i.e. accelerometers, to explore features to analyse FoG episodes. They found that the frequency response in 2 patients wearing accelerometers at the ankle was around 6–8 Hz [19]. In 2008, Moore et al. proposed a methodology for detecting FoG with an accelerometer at the ankle where they described the Freezing Index (FI), which is the quotient of the power spectral density (PSD) from 3 to 8 Hz (Freezing Band, FB) and the PSD from 0.5 to 3 Hz (Walking Band, WB) [21]. When the FI exceeds an identified threshold (Freezing Threshold, FTH), a FoG episode is considered to have occurred. Due to the presence of false positives (FP) when the patient is at rest, Bächlin et al., in 2009, introduced the Power Index (PI), defined as the addition of WB and FB, which was compared to the Power Threshold (PTH) to establish if there was any relevant quantity of movement when FI was high [21]. PI indicates the amount of movement, therefore, situations where patients were not moving voluntarily were eliminated. In this final algorithm, a FoG episode is then determined to occur if  $FI > FTH$  and  $PI > PTH$ . The MBFA algorithm is the most advanced FoG detection algorithm due to its low computational cost and good performance [22–24,26,27,31,46–50].

The MBFA algorithm has been widely employed to analyse FoG, although usually in laboratory conditions and very often with low patient numbers. Jovanov et al. implemented this algorithm in real time, although only one patient was used to test the algorithm. Furthermore, no results on sensitivity and specificity were reported [22]. Zabaleta et al. analysed FoG by means of a triaxial accelerometer and a biaxial gyroscope in different locations of the lower limbs. The main features employed were the FI in conjunction with power spectral densities changes. They were capable of correctly detecting 82.7% of FoG episodes with inertial sensors located at each ankle, though only in 2 PD patients [24]. Other attempts to analyse FoG include Stamatakis et al. [51] and Takac et al. [44], but they only used one patient to test the system for PD gait extraction and a system for contextualising FoG events, respectively.

In recent years, Niazmand et al. (2011) presented the Mimed-pants [26], a washable jogging pants with five integrated accelerometers. They employed MBFA for detecting FoG achieving an 88.3% on sensitivity and 85.3% on specificity in 6 PD patients in a short and controlled test focused on inducing FoG regardless of a false positives test. In a recent work in 2012, Zhao et al. [46] performed an embedded algorithm based in the MBFA approach within the Mimed-pants system obtaining 81% for sensitivity with 8 PD patients using similar tests. More recently, Mazilu et al. proposed new online algorithms using 3 accelerometers and comparing different machine learning classifiers that exploit MBFA features, as well as few other ones, in

10 PD patients [48]. Results obtained showed higher than 95% for specificity and sensitivity with different classifiers. However, tests were performed under controlled conditions and, furthermore, the validation methodology overestimated the performance measures since classifiers were trained, iteratively, with all window signals available from a patient except for one window, which was used to obtain the performance values. Therefore, training and testing patterns were very similar, which is very different from normal ADLs. Thus, the reported specificities and sensitivities would be expected to dramatically decrease free-living conditions. In 2013, Moore et al. published their most recent work focused on the MBFA. In this work, they compared different configurations applying the same algorithm on 25 PD patients, among whom 20 had FoG episodes. Different window sizes, sensor locations, and different values for FTH and PTH were evaluated in order to find the optimal conditions. Results showed better performances with longer windows; however, with a long window size, Moore et al. reported a relevant loss of sensitivity at short episodes that, paradoxically, are the most frequent in PD patients [27]. In a more complex test than the one performed previously [20] using up to 7 sensors and a longer test protocol, they achieved sensitivity and specificity over 70% but, in some configurations (window size: 7.5s) with the system worn on the back of the waist, both sensitivities and specificities were over 80%. In a different approach, Tripoliti et al. proved different features and locations in order to also know the best configuration [52]. The work was performed with 5 PD patients and under controlled conditions, with a specific protocol designed to elicit FoG and without FP protocol tests. In this work, they also used 2 gyroscopes along with 6 accelerometers located in different body parts. With all the sensors worn, they obtained an accuracy of 96.11%, 98.74% of specificity and 81.94% of sensitivity with a leave-one-patient-out method. On the other hand, with a single IMU on the waist they report a 75% of sensitivity and 95% of specificity, although the algorithm is not compared with any other method under the same conditions.

In a more recent study from Mazilu's group, they investigated unsupervised feature learning for building an optimal input vector for a decision tree classifier with the DAPHNET project dataset (10 PD patients). Their approach was compared to an analogous MBFA method in which the FI and the energy of the spectral band between 0.5Hz and 8Hz were evaluated. The training and testing was user-dependent and under controlled conditions. Results enhanced the similar MBFA method by up to 8.1% in terms of F1-score. Another approach was presented by Rodríguez et al. who proposed a method to contextualise FoG episodes by means of an activity recognition algorithm, which rejected those false positives when the patient was sitting and performed activities such as drawing or typing on a laptop. The specificity was improved by up to 5% on average reaching an enhancement of 11.9% in some cases [50]. However, this added-contextualisation method did not contribute to enhanced sensitivity. Other works have studied the variability of gait between a FoG episode and normal gait. Although the results are interesting, they failed to include false positive results and a reliable classification was not performed [53,54]. A recent paper by Zach et al. presented a new methodology to elicit FoG under laboratory conditions which was evaluated with the MBFA algorithm obtaining 75% on sensitivity and 76% on specificity [31].

Finally, Ahlrichs et al., within the REMPARK project [55], used Support Vector Machines (SVM) for detecting FoG episodes in 8 patients with PD at their own homes. The method included tests in different motor states and used a single accelerometer at the waist, achieving accuracies of over 90%. However, specificity was only computed with non-FoG patients, which could lead to unreliable predictions since the models were not tested with PD patients with FoG who move slightly differently to those patients without FoG [30]. Furthermore, the evaluation was performed over a single minute, which is considered too long if cueing is desirable [28].



Tables 1 and 2 summarise the methods presented in the work we have highlighted.

## Methods

This section describes the new approach proposed by the authors to detect FoG with a supervised machine learning technique and the experimental methods used to evaluate it. The section is divided into 4 sub-sections, (Fig 1) that describes the steps performed in the methodology. At the end of the section, there is a final paragraph devoted to the tests performed to evaluate the MBFA algorithm. The first sub-section describes the database collection procedure, the protocol tests and baseline data from patients. The second sub-section reports the signal processing methods for the SVM model and the extraction of features of the signals based on the online implementation requirements. The third section details the procedure to derive an SVM-based classifier based on a generic model for all patients and in a personalised model. Finally, the evaluation for FoG episodes is described in the fourth section, with respect to the personalised and the generic approaches. In addition, a fifth sub-section has been included, in which the assessment of the MBFA is also explained with the aim to compare the proposed approach to this widely extended algorithm. Although the evaluation of the classifier is the same as the proposed model, the training and feature extraction processes are different; this is detailed in this final section.

## Data collection

The study was approved by local Ethics Commission: Spanish Agency for Drugs and Medical Devices (Spain), Clinical Research Ethics Committee-Galway University Hospitals, Health Department Maccabi Healthcare Services and Ethics Committee Fondazione Santa Lucia. PD patients were diagnosed according the UK Brain Bank [60] criteria and all of them gave their written informed consent to perform the datasets.

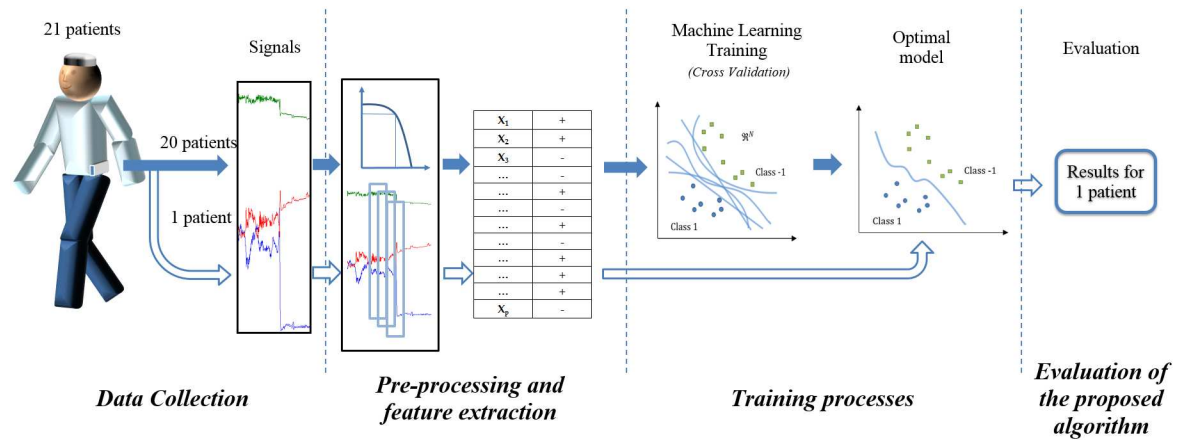
A total of 21 patients (3 women and 18 men) with PD participated in the experimental protocol to collect signals. The experimental protocol was part of the REMPARK project [55], which grouped a total of 92 patients with PD (36 women, 56 men) from 4 health centres in Europe, Centro Médico Teknon (Spain), Maccabi Healthcare Services (Israel), Fondazione Santa Lucia (Italy) and National University of Galway (Ireland) [61].

In REMPARK's experimental protocol, all patients performed the tests both in "OFF" and "ON" states. "OFF" states are those periods where rigidity, bradykinesia or postural alterations appear in PD patients. During the "ON" state, these symptoms are mitigated due to medication effect although other side effects may appear, such as dyskinesias. The inclusion criteria in this clinical protocol were Hoehn & Yahr stage [62] greater than 2 in the OFF state and the absence of dementia according to DSM IV criteria [60]. In this work, which is focused on the detection of FoG episodes, some additional inclusion criteria have been required. First, PD patients should have a FoG Questionnaire (FoG-Q) [15] score above 6, and only those PD patients whose total FoG duration was at least one minute long were selected. Overall, 21 patients out of the 92 patients matched all inclusion criteria for this study (Table 3). The mean age was  $69.29 \pm 9.72$  years old and the mean Hoehn & Yahr stage scoring was  $3.07 \pm 0.43$  in the OFF state. Sixteen patients needed walking assistance and the mean FoG-Q test index was  $15.8 \pm 4.11$ , the lowest score being 7 and the highest score 23. Six out of 21 patients did not exhibit dyskinesias, and eleven patients had wearing-off phenomena. At time of the tests, all patients were medicated with optimal doses of levodopa and dopamine agonist therapy. Some baseline data from patient 5 were lost, which is shown as N/A in Table 3; however, they do not affect to the results of this study.

**Table 2. Summary of published works regarding FoG detection algorithms.**

<i>Work</i>	<i>Year</i>	<i>Algorithm methods</i>	<i>Main significant results presented</i>
Nieuwboer et al. [39]	2001	Objective measurements and contrast	Reduction of stride length, increase of step cadence compared to normal gait
Han et al. [19]	2003	Statistical test	Observation of frequency response from 6-8Hz in a FoG episode
Hausdorff et al. [38]	2003	Statistical test	Observation of frequency response in FoG from 3-6Hz compared to normal walking
Moore et al. [20]	2008	Threshold based	General threshold: 78.3% success; Personalised threshold: 89.1% success
Daphnet project [56]	2009–2011	Threshold based	General threshold.: 73.1% Sensitivity, 81.6% Specificity; Personalised threshold: 88.6% Sensitivity, 92.8% Specificity
Zabaleta et al. [24]	2008	Linear Classifier through FI variable	82.7% success
Jovanov et al. [22]	2009	Threshold based	Algorithm in real time with fast response. Results of algorithm performance not provided
Maidan et al. [57]	2010	Statistical test	Observation of heart rate evolution during FoG episodes
Knobl et al. [40]	2010	Statistical test	Observation of main differences of a normal gait and gait with FoG
Delval et al. [23]	2010	Threshold based	75–83% Sensitivity, >95% Specificity
Cole et al. [25]	2011	Dynamic Neural Networks	83% sensitivity 97% specificity
Niazmand et al. [26]	2011	Threshold based	88.3% Sensitivity and 85.3% Specificity
Stamatakis et al. [51]	2011	Statistical test	Main differences observation of a normal gait and gait with FoG
Almeida et al. [40]	2012	Statistical test	Gait parameters observation with cues between healthy people and PD patients
Zhao et al. [46]	2012	Threshold based	81.7% Sensitivity, Specificity not provided
Handojoseno et al. [42]	2012	Neural networks through wavelet analysis	75% Accuracy, 75% Sensitivity, 75% Specificity
Mancini et al. [47]	2012	Statistical test	Analysis of frequency ratio and distinguishing objectively PD patients with and without FoG and healthy people
Mazilu et al. [48]	2012	Different classifiers and different temporal-frequency features	Online machine learning system with >95% sensitivity and >95% specificity with some configurations
Takac et al. [58]	2013	Threshold based for the FoG algorithm, Artificial Neural Networks for context algorithms	17 degrees and 0.16 m error (RMSE) for human body orientation
Moore et al. [27]	2013	Threshold based	Contrast of different configurations (window size, sensor location, freezing and power thresholds) Sensitivity and Specificity >70%
Tripoliti et al. [52]	2013	Different classifiers using entropy	96.11% Accuracy
Mazilu et al. [49]	2013	Feature learning with decision tree classifier	Better performance compared to classical methods, approach for detecting pre-FoG
Rodríguez et al. [50]	2014	Threshold based + machine learning context based algorithm	Improvement of 5% in Specificity due to posture contextualisation
Coste et al. [53]	2014	Threshold based through gait parameters	Approach to detect pre-FoG based on gait parameters
Weiss et al. [30]	2015	Statistical test	Observation of gait parameters between freezers and non-freezers
Tay et al. [54]	2015	Threshold based for extracting gait parameters	Observation of variability in gait parameters
Zach et al. [31]	2015	Threshold based	75% Sensitivity, 76% Specificity
Mazilu et al. [43]	2015	Threshold based through Gaussian kernel	Pre-FoG with 71.3% of success
Ahlich et al. [28]	2015	RBF kernel function through Support Vector Machines	Accuracies >90%
Maidan et al. [59]	2015	Objective measurements	Observation of variability in levels of haemoglobin just before a FoG episode

doi:10.1371/journal.pone.0171764.t002



**Fig 1. Methods description.**

doi:10.1371/journal.pone.0171764.g001

All patients performed two tests, each approximately 20 minutes long, which comprised a set of scripted activities. All activities were carried out at the patients’ homes, where the probability of eliciting FoG episodes was expected to increase [17]. Medication effect has a strong influence regarding the motor response in PD patients with FoG [11,14,63], thus, tests were performed with and without medication. The first test took place early in the morning before patients had taken their first medication and they felt in OFF state. Thus in the first test, patients completed the different activities without antiparkinsonian medication. Once patients had completed the different activities without medication, they took their usual medication

**Table 3. Clinical characteristics year 2013.**

Patient's index	Gender	Age	Year of Diagnosis	Hoehn & Yahr (OFF)	UPDRS part III (OFF)	UPDRS part III (ON)	FoG Questionnaire	MMSE
1	M	83	2008	3	28	14	12	30
2	M	43	2008	2.5	42	5	11	29
3	M	66	2004	4	38	8	17	30
4	M	74	2002	4	53	35	18	25
5	M	72	N/A	3	N/A	N/A	N/A	N/A
6	M	73	2002	3	59	36	17	27
7	M	74	1995	3	53	28	20	26
8	F	69	1997	3	56	15	20	29
9	M	79	1999	3	36	16	12	26
10	M	77	1999	3	43	22	13	25
11	M	72	2002	3	44	29	23	29
12	M	50	2006	2.5	48	9	7	30
13	M	80	2004	3	50	19	23	28
14	M	60	1998	2.5	23	5	17	29
15	F	74	2000	4	33	18	19	27
16	M	60	2001	3	12	3	16	29
17	M	73	2005	3	20	12	16	27
18	M	69	2003	3	23	10	14	29
19	M	74	2008	3	15	11	12	30
20	M	62	2009	3	23	18	13	24
21	F	71	2004	3	27	10	16	26

doi:10.1371/journal.pone.0171764.t003



**Fig 2. The 9x2 and location of the IMU.**

doi:10.1371/journal.pone.0171764.g002

and, at least one hour later, the second test was performed (i.e. ON state). Although the protocol test consisted of specific scripted activities, the test also allowed patients to perform several activities in a free moving manner. Furthermore, the activities were performed in a familiar environment and the patient could proceed with their normal ADLs in a natural way. More specifically, the activities performed in both tests consisted of I) showing the researchers around their home; II) Stand Up and Go test crossing through a doorway and then, turning back (FoG provocation test), which was repeated up to 10 times; III) going outdoors and taking a short walk; and IV) a dual task activity (e.g. reading while carrying an object). In the second test, i.e. following medication, a false positive protocol activity for FoG was additionally performed. This protocol test relies on performing different activities in which the patient executes a short and fast movement repeatedly whose inertial frequency content is similar to a FoG episode. The executed activities were brushing their teeth, painting/drawing/erasing on a sheet of paper and cleaning windows.

A total of 93.03 minutes of FoG episodes were registered (1321 episodes) among these 21 patients, among which 74.12 minutes (79.67%) were registered when patients were in OFF state, 12.35 minutes (13.28%) were registered when patients were in ON state and 6.56 minutes (7.05%) were registered in an intermediate state.

Acceleration signals were captured while patients executed the described protocol with the 9x2 Inertial Measurement Unit (IMU), a small ( $77 \times 37 \times 21 \text{ mm}^3$ ) and light (78g with battery) wearable inertial system (Fig 2) which was located at the left side of the waist [64]. This IMU, whose autonomy lasts more than a day, stored the accelerometer data at 200Hz in a microSD card.

All tests were video-recorded and videos were synchronised with the signals recorded by the 9x2 device through a specific method [37]. In order to create a gold standard for the occurrence of FoG episodes, videos were labelled offline by an experienced clinician. Within the MASPARK project [36], all activities, postures and partial FoG episodes were labelled by clinicians.

### Pre-processing and feature extraction

This subsection presents a description of the employed method for pre-processing the signal, which comprises signal conditioning, windowing and feature extraction. Note that, since the proposed algorithm is thought to be implemented in the near future within the 9x2 device for its real-time execution, some computational issues have been taken into account to facilitate this future implementation. For example, windows with reduced amount of samples are considered in order to, first, reduce the latency time of the algorithm; second, to deal with the

Number of features	Feature Group	Feature	Equation
5	1	Means	$\bar{x}^W = \frac{1}{N} \sum_{i=1}^N x_i^W, \bar{y}^W = \frac{1}{N} \sum_{i=1}^N y_i^W, \bar{z}^W = \frac{1}{N} \sum_{i=1}^N z_i^W, \bar{xz}^W = \bar{x}^W - \bar{z}^W, \bar{yz}^W = \bar{y}^W - \bar{z}^W$
1	2	Increments	$\bar{x}^W = \bar{x}^W - \bar{x}^{W-1}$
1	3	Incr.(mean(Y)-mean(X))	$\bar{yx}^W = (y^{iW} - x^{iW}) - (y^{iW-1} - x^{iW-1})$
1	3	Incr.(mean(X)-mean(Z))	$\bar{zx}^W = (x^{iW} - z^{iW}) - (x^{iW-1} - z^{iW-1})$
3	4	std(X, Y, Z)	$\sigma_{x^W} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i^W - \bar{x}^W)^2}$
3	5	Corr (X,Y), Corr(X,Z), Corr(YZ)	$\rho_{y,x}^W = \frac{\sum_{i=1}^N ((y_i^W - \bar{y}^W)(x_i^W - \bar{x}^W))}{N \sigma_{y^W} \sigma_{x^W}}, \rho_{z,y}^W = \frac{\sum_{i=1}^N ((z_i^W - \bar{z}^W)(y_i^W - \bar{y}^W))}{N \sigma_{z^W} \sigma_{y^W}}, \rho_{z,x}^W = \frac{\sum_{i=1}^N ((z_i^W - \bar{z}^W)(x_i^W - \bar{x}^W))}{N \sigma_{z^W} \sigma_{x^W}}$
5	6	Freq. std (Y)	$\sigma( V_{0.1,0.68}^W ), \sigma( V_{0.68,3}^W ), \sigma( V_{3,8}^W ), \sigma( V_{8,20}^W ), \sigma( V_{20,8}^W )$
3	7	Higher harmonics (Y)	$\max( V_{0.1,20}^W ), 2ndmax( V_{0.1,20}^W ), f_{\max}( V_{0.1,20}^W ) - f_{2ndmax}( V_{0.1,20}^W )$
1	8	Com (Y)	$\frac{\sum_{i=1}^N ( V_{i-1}^W  \cdot f_{ V_{i-1}^W })}{\sum_{i=1}^N  V_{i-1}^W }$
7	9	Skewness (X,Y,Z)	$\gamma_{x^W}^W = E \left[ \left( \frac{x^W - \bar{x}^W}{\sigma_{x^W}} \right)^3 \right], \gamma_{y^W}^W = E \left[ \left( \frac{y^W - \bar{y}^W}{\sigma_{y^W}} \right)^3 \right], \gamma_{z^W}^W = E \left[ \left( \frac{z^W - \bar{z}^W}{\sigma_{z^W}} \right)^3 \right], \gamma_{yx}^W = E \left[ \left( \frac{(y^W - \bar{y}^W)(x^W - \bar{x}^W)}{\sigma_{yx}^W} \right)^3 \right]$ $\gamma_{yz}^W = E \left[ \left( \frac{(y^W - \bar{y}^W)(z^W - \bar{z}^W)}{\sigma( V_{0.68,3}^W )} \right)^3 \right], \gamma_{zx}^W = E \left[ \left( \frac{(z^W - \bar{z}^W)(x^W - \bar{x}^W)}{\sigma( V_{0.68,3}^W )} \right)^3 \right], \gamma_{yz}^W = E \left[ \left( \frac{(y^W - \bar{y}^W)(z^W - \bar{z}^W)}{\sigma( V_{0.68,3}^W )} \right)^3 \right]$
7	9	Kurtosis	$\gamma_{x^W}^W = \frac{E[(x^W - \bar{x}^W)^4]}{(E[(x^W - \bar{x}^W)^2])^2}, \gamma_{y^W}^W = \frac{E[(y^W - \bar{y}^W)^4]}{(E[(y^W - \bar{y}^W)^2])^2}, \gamma_{z^W}^W = \frac{E[(z^W - \bar{z}^W)^4]}{(E[(z^W - \bar{z}^W)^2])^2}, \gamma_{yx}^W = \frac{E[(y^W - \bar{y}^W)(x^W - \bar{x}^W)^2]}{(E[(y^W - \bar{y}^W)^2]) (E[(x^W - \bar{x}^W)^2])},$ $\gamma_{yz}^W = \frac{E[(y^W - \bar{y}^W)(z^W - \bar{z}^W)^2]}{(E[(y^W - \bar{y}^W)^2]) (E[(z^W - \bar{z}^W)^2])}, \gamma_{zx}^W = \frac{E[(z^W - \bar{z}^W)(x^W - \bar{x}^W)^2]}{(E[(z^W - \bar{z}^W)^2]) (E[(x^W - \bar{x}^W)^2])},$
3	10	PCA	$PCA( S_N^W )$
3	11	Integrals	$\int_{i=1}^N x_i^W, \int_{i=1}^N y_i^W, \int_{i=1}^N z_i^W$
12	12	Auto regression coefficients	$ac_x^W = arburg(x_N^W, 4), ac_y^W = arburg(y_N^W, 4), ac_z^W = arburg(z_N^W, 4), a \in \mathbb{R}^4$

Fig 3. Set of features employed as input vector for the SVM classifier.

doi:10.1371/journal.pone.0171764.g003

existing memory restrictions of current microcontrollers, which cannot store inertial information during large periods of time; and third, to prevent the feature extraction process from becoming a heavy computational burden.

The 9x2 provides a 200Hz inertial signal in each of the 3 axis of the accelerometer. This data is resampled at 40Hz since this frequency has been shown to be enough to analyse human movements [65] and it is sufficient to analyse the frequencies obtained during FoG episodes [20]. Once resampled, inertial signals are filtered in order to remove the high frequency noise through a 2<sup>nd</sup> order Butterworth low-pass filter with a 15Hz cut-off frequency.

The next step is to window the signal. According to Moore and colleagues, the specificity in very small window sizes is very low, thus, it is recommended to work with windows over 2.5 s [27]. In other previous work, Bächlin et al. worked in window sizes of 4s [29], Mazilu et al. showed how FoG detection performance was kept with windows over 3s [48] and Zach et al. showed that window lengths between 2 and 4 s. provided similar results [31]. In order to achieve a similar window length and given that the sampling frequency is 40Hz, a window length of 128 samples (N = 128), i.e. 3.2s of window-length, is proposed. Windows are overlapped at 50% so as to prevent information loss between windows.

With the aim to obtain significant information from each window, some features are extracted. These features will be the input vector,  $x_i$ , for the proposed classifier in *Training processes* section, which is based on machine learning techniques. In this work, a total of 55 features are employed, which are presented in Fig 3. Features are briefly described below by grouping them according to their purpose. The number of feature group is noted within brackets.

1. Mean value of each accelerometer axis measurements throughout the window gives the orientation of the inertial system related to gravity in absence of movement.
2. Increments of consecutive windows' mean values in each accelerometer axis provide the amount of variation performed in each axis of the accelerometer.
3. Difference between the increments of the windows' mean values among different axes gives the amount of variation among different axes. This feature is significant regarding the detection of postural transitions [66].
4. Standard deviation of each axis indicates the amount of movement performed in a window time.
5. Correlation among each pair of axes gives information about linear relation in different axes, which is useful for detecting normal/continuous gait.
6. Standard deviation on specific spectral bands provides information about the distribution of harmonics in these bands, which define if the harmonics are distributed along the band or if there are significant harmonics. It can give information about stability or specific significant peaks in a band.
7. Highest harmonic peaks are useful for detecting concrete events in specific spectral bands (e.g. high peaks in 0.1Hz to 0.68 Hz band can be considered as a candidate to be a postural transition) [66–68].
8. Spectral density centre of mass gives information about the overall quantity of movement performed and the frequency band in which most of the movement is concentrated.
9. Skewness and kurtosis are employed to characterise the spectral distribution.
10. A change of basis, previously obtained by Principal Component Analysis (PCA), is performed over the energy in the spectral band harmonics from 0.1Hz to 8 Hz. The aim of this is to characterise the harmonics contained in the FB and WB band by features that maximise the variance in this band.
11. Integrals of the accelerometer axis gives information related to the quantity of movement.
12. Auto regression coefficients provide information of how the signal is correlated with a time-shifted version of itself giving the amount of time shift [69].

In order to clarify some feature notation reported in Fig 3, see clarifications below. Let  $x_1^W, \dots, x_N^W$  be a set of samples in a certain window  $W$ . The Short Time Fourier Transform (STFT) is composed of complex values  $X_1^W, \dots, X_N^W$  obtained following the next equation:

$$X_h^W = \sum_{n=1}^N x_n e^{-i2\pi h \frac{n}{N}} \tag{1}$$

where  $h = 1, \dots, N$ . The output of the STFT ( $X_h^W$ ) is a set of  $N$  complex numbers, each one of them representing the amplitude and phase of a harmonic. Let  $|X_{f_1, f_2}^W| = \sum_{h=f_1}^{f_2} |X_h^W|$  be, hence, a set of values representing the sum of the absolute value of  $X_h^W$ , where  $f_1$  and  $f_2$  are the upper and lower frequency harmonics denoting the set of harmonics establishing the band to be analysed in the current  $W$  window. This parameter is employed in the features group 6 to 10. Otherwise, let  $M_i^W = \sqrt{x_i^2 + y_i^2 + z_i^2}$  be the modulus of the 3 accelerometer axes in a sample  $i$  and a window  $W$ . In addition, the highest and the second highest harmonic within the frequency

spectrum analysed are denoted as  $\max(A)$  and  $2\text{ndmax}(A)$  where  $A = [A_1^W, \dots, A_N^W]$  is the set of harmonics from a signal axis (feature group 7).

Finally, the Principal Component Analysis (PCA) has also been employed, as described below. Let  $|S_h^W| = |X_h^W| + |Y_h^W| + |Z_h^W|$  be the summation of harmonics for each accelerometer axis, where  $h = 0.1, \dots, 8$  Hz, so that a vector  $S^W = [S_1^W, \dots, S_N^W]$  is obtained, which represents the spectral information in the [0.1, 8] Hz band in a certain window. PCA has been used to reduce the dimensionality of these harmonics. More specifically, PCA has been applied to a set of  $S^W$  vectors organised in a matrix  $M = [S^{1T}, S^{2T}, \dots, S^{LT}]^T$ . These  $L$  windows are a randomly selected subset of the training data composed by a 10% of the data from each patient, thus containing both FoG episodes and other signals. PCA is applied in order to obtain the orthonormal change of basis matrix  $V$  that brings each pattern  $S^W$  from the original space to the latent one. PCA is applied to matrix  $M$  by first centring the data, and then applying a Singular Value Decomposition (SVD) so that  $M = U\Sigma V$  is obtained. In the SVD process, it was identified that only the first three latent variables contributed with data (i.e. their singular values were significantly higher than the rest). In consequence, only the first three components of the latent representation of  $S^W$ , i.e.  $V S^W$ , are employed as features (feature group 10). Note that the PCA process is only applied once in the training process and a real-time implementation of the feature-extraction process would only require the implementation of the product  $V S^W$ .

## Training processes

The symptoms related to FoG episodes are user dependent with typical symptoms manifesting as gait abnormalities (e.g. shuffling, leg trembling in place and akinetic episodes) [11]. As each FoG patient experiences these symptoms in different ways, developing an algorithm to detect FoG episodes from inertial signals can be challenging due to the variation between individual signals. To overcome these challenges previous researchers have developed user dependent algorithms capable of detecting FoG episodes with a high degree of accuracy [21,48]. These algorithms require the prior capturing of the patients inertial signals in order to train the algorithm, thus making the algorithm cumbersome to obtain. A less challenging solution would be to develop a generalised algorithm, since it would only require the obtaining of inertial signals from a relatively small cohort of FoG patients in order to train the algorithm for use with any patient. To investigate both types of algorithms and their associated training processes we proposed two methodologies I) a generic model through a leave one out method and II) a model that is personalised to each patient.

**Generic FoG model selection.** In order to obtain a generic (user independent) model, a leave-one-patient-out methodology has been adopted. The model is generic in the sense that it is not personalised to any patient and there is no adjustment made to the model that considers the patient for which the model is tested on. Given the nature of the problem at hand, Support Vector Machines (SVM) have been considered to be the most suitable classifier, due to their high performance method for two-class classification and ease of implementation within microcontrollers [70]. SVM is a machine learning technique that use the *kernel trick* to map input vectors ( $x_i \in \mathbb{R}^n$ ,  $1 \leq i \leq p$ ) to a higher dimensional feature space by  $\varphi(x_i)$ , where typically  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$  and subjected by  $m > n$ . The kernel trick enables SVM to avoid bringing  $x_i$  to the higher dimensional space by computing the dot product of patterns within  $\mathbb{R}^n$ .

In our approach, each  $x_i$  is represented by 55 parameters (previously described in the *Pre-processing and feature extraction* section) which characterise the current window  $W$ . The input vectors are associated to one of two different classes, FoG or no-FoG. These classes are represented by the labels  $y_i = \{1, -1\}$ ,  $1 \leq i \leq p$ , where  $y_i = 1$  in case of a FoG pattern and  $y_i = -1$  in case of a non-FoG pattern. The set of  $y_i$  values is provided by clinical experts with the help of

video recordings that were synchronised with their associated inertial signals. In order to maximise the margin, reduce the empirical error and minimise the structural risk, a hyperplane is built to separate the 2 classes [71]. The optimal hyperplane is defined according to the following optimisation problem:

$$\begin{aligned} \text{minimize } W(\alpha) &= \sum_{i=1}^p \alpha_i - \frac{1}{2} \sum_{i,j=1}^p \left( \alpha_i \alpha_j y_i y_j \cdot K(\mathbf{x}_i, \mathbf{x}_j) \right) \\ \text{s. t.} \quad & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, p \end{aligned} \tag{2}$$

in which  $C$  is the trade-off between the classification margin and misclassification error, and the kernel function  $K(\mathbf{x}_i, \mathbf{x}_j)$  replaces the dot product  $\varphi(\mathbf{x}_i) \cdot \varphi(\mathbf{x}_j)$ . The kernel used for the FoG classifier is a Gaussian radial basis function (RBF), due to its good performance and generalisation capacity,  $K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2}$ , where  $\gamma$  is the width of RBF function.

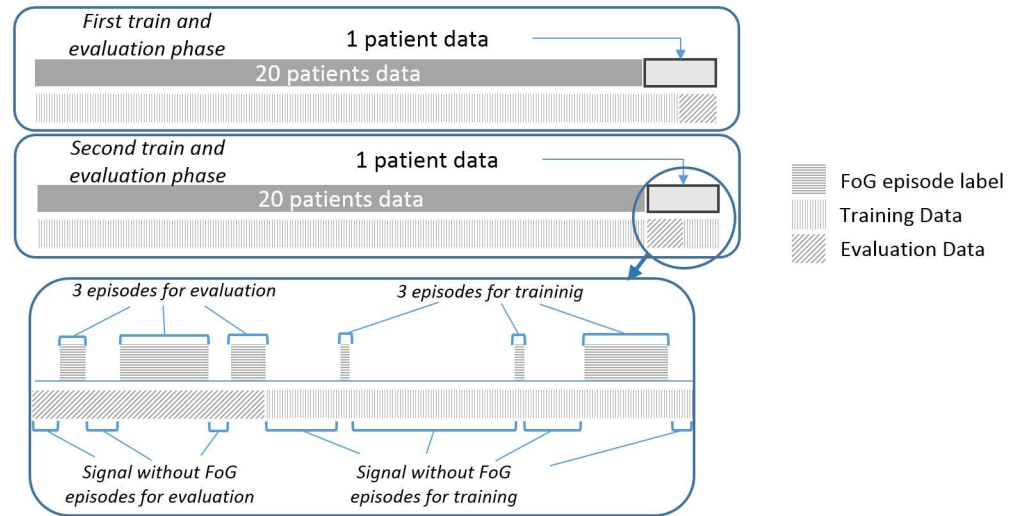
Due to the leave-one-patient-out approach, these parameters must be adjusted with the data from all patients except one. Thus, with this data, a range of values  $\{10^{-2}, 10^{-1}, \dots, 10^2\}$  for  $C$  and  $\gamma$  are tested through a 10-fold cross-validation method. From the results of the different folds the average sensitivity and specificity is obtained for each combination of values. Optimal parameters for the SVM ( $C_{opt}, \gamma_{opt}$ ) are those that maximise the following expression:

$$\begin{aligned} (C_{opt}, \gamma_{opt}) &= \operatorname{argmax}(\sqrt{\text{Sensitivity}_{C,\gamma} \cdot \text{Specificity}_{C,\gamma}}) \\ \text{s. t. } & \text{Sensitivity}_{C,\gamma} > 70\%, \text{ Specificity}_{C,\gamma} > 70\% \end{aligned} \tag{3}$$

Once found, the optimal values are used to train the SVM model. The data from the patient that is still unused is then employed to test the model with a specific procedure, which is described in the *Evaluation of the proposed algorithm* section. The entire training process is repeated by keeping apart the data from each patient; in consequence, the number of SVM models obtained is the same as the number of available patients.

**Personalised FoG model selection.** In this section, the proposed user-dependent method is described. In order to obtain a user-dependent model, a modified leave-one-patient-out method has been adopted, which also has been performed by means of SVM. In contrast to the generic model presented in the previous subsection, the personalised model requires labelled inertial signal data from the user in order to obtain the personalised classifier before using it in real life. Therefore, we will use a portion of the data from the patient being tested to train our classifier, and the remaining portion of data will be used to evaluate the algorithm. Signals from the patient being tested are divided into two sets, the first set being composed of 50% of the labelled FoG episodes with the corresponding non-episodic signals, and the second set comprising the remaining FoG episodes and non-episodic windows. Only the first set, i.e. the first half, is included in the training dataset, as shown in Fig 4. We note that the optimal choice of the training data would be to include 50% of the signals into the training dataset (without splitting by episodes) and the remaining 50% into the evaluation dataset. However, there are patients who mostly freeze during the initial data collection while other patients only presented FoG during the second part of the data collection. Therefore, the number of episodes in each dataset would be, in most cases, unbalanced. Hence, the training dataset is chosen to be those parts of the signal from the testing patient that comprise 50% of the FoG episodes, together with the non-episodic signals, as shown in Fig 4. Due to these requirements and the evaluation process, episodes cannot be selected randomly. If FoG episodes were selected randomly, those parts of the signal that do not contain FoG episodes would have to be randomly selected as





**Fig 4. Example of training and evaluation data used in the personalised FoG model selection.**

doi:10.1371/journal.pone.0171764.g004

well. Consequently, the proposed evaluation method would not be valid, since, as shown in the next Section, True Positives (TP) and False Negatives (FN) depend on episode durations, and True Negatives (TN) and False Positives (FP) depend on the amount of detected time spent without a positive detection. Thus, random episode selection would make detection's temporal line disappear and evaluation by episode would be unfeasible.

Likewise, the same analysis is performed again but swapping the training and evaluation datasets of the testing patient employed to train and to validate the model. In consequence, after first training the model with data from non-testing patients and the first set of the data from the testing patient, a second training takes place. In this case, the second set comprising of the last 50% of the episodes, including non-episodic signals, of the testing patient are included into the training dataset as shown in the lower part of Fig 4. The first 50% of the episodes and its corresponding non-episodic signals are then employed for evaluating the obtained model. Although other percentages of training/evaluation could be tested, in this case, the authors consider that the 50% of FoG episodes' data contains enough information of non-episodic signals as well.

In this modified leave-one-out approach, data from the non-testing 20 patients represent a clearly greater dataset compared to data from the patient being tested; thus, data from the latter are almost negligible and they would not have any impact in the training process. Hence, it is considered interesting to force the model selection process to learn from the testing patient's data with more influence than the remaining patients. More specifically, the weight of the testing patient's data is set to be the same as the data from the other 20 patients. Thus, data belonging to the testing patient, represented by  $X = \{x_i\}$  is weighted in the following form. Let  $n_{20}$  be the number of patterns of the data belonging to 20 patients used to train the model. On the other hand, let the data from the testing patient be divided into two parts with the first and the second 50% of FoG episodes as previously described, each part containing  $n_f$  and  $n_s$  patterns respectively. The weight given to the those patterns from the training dataset corresponding to the testing patient, when the first half is used, is  $w_t = \frac{n_{20}}{n_{20} + n_f}$ . On the other hand, the weight given to those patterns from the training dataset belonging to the 20-patients' data is  $w_{20} = \frac{n_f}{n_{20} + n_f}$ . Accordingly, for the second half,  $n_s$  would be used instead of  $n_f$ . Thus, the weight given to each part corresponds to the prior of the alternative part of the training dataset. These

weights are included in the SVM model according to the following equation:

$$\text{maximize } W(\alpha) = \sum_{i=1}^p \alpha_i - \frac{1}{2} \sum_{i,j=1}^p (\alpha_i \alpha_j y_i y_j \cdot K(x_i, x_j)) \quad (4)$$

$$\text{s. t. } 0 \leq \alpha_i \leq C \cdot w_{20}, \quad x_i \notin {}^tX$$

$$0 \leq \alpha_i \leq C \cdot w_t, \quad x_i \in {}^tX$$

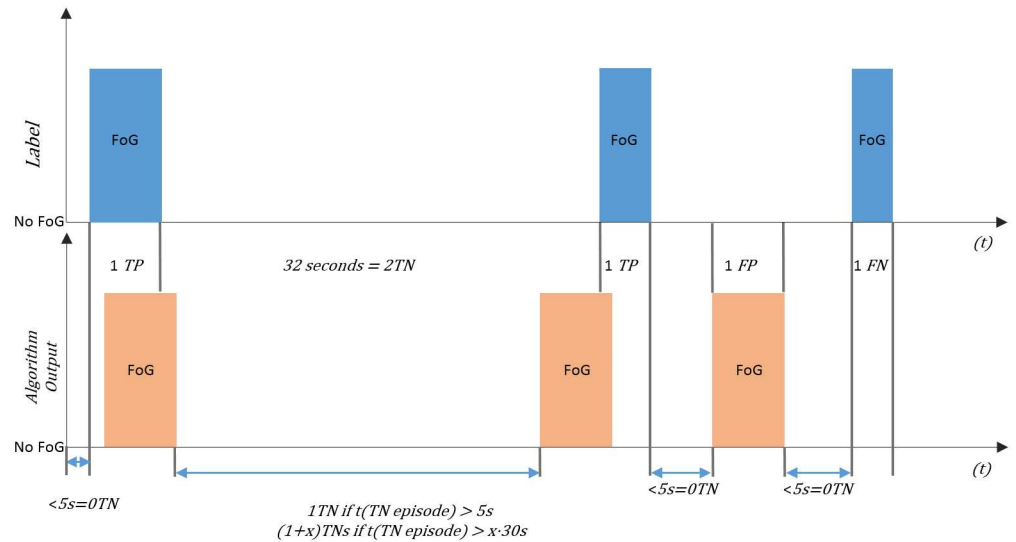
Finally, the personalised model obtained through the training process will be evaluated over the other set of windows which have not been used in the training according to the episode-based evaluation previously presented. Afterwards, the methodology for obtaining the SVM model will be repeated switching the parts used in the training and evaluation. Thus, the model is prevented from overfitting as only one part of the data is performed by the patient, which may coincide with a specific test and not with ADL. Finally, to ensure each patient's data is used once as a testing patient the process is completed 21 times.

## Evaluation of the proposed algorithm

To analysis, the performance of each methodology, a rigorous evaluation is performed which follows the methodology previously described [72]. Due to likely situations in which patients are sitting, standing, or lying for long periods of time the proportion of TN (conditions in which a FoG episode was not detected by the algorithm and the patient was not in a FoG episode) becomes extremely large. Similarly, due to situations in which patients experience long FoG episodes the proportion of TP (conditions in which a FoG episode was detected by the algorithm and the patient was in a FoG episode) also becomes extremely large. The purposed evaluation process aims to balance the excessive number of TN and TP that may appear with window-based evaluations of machine learning approaches through the evaluation analysis of episodes instead of the evaluation analysis of windows that contain an episode. Therefore, a FoG episode is evaluated from its beginning until its end as a single episode, instead of being evaluated within every window that contains the FoG episode.

Given that during the data collection protocol the minimum time taken by patients to perform the false positive test was 38s, it was considered that the maximum time of a single TN episode should be 30 seconds. Thus, for example, if a patient is executing some activity without any FoG episode during 40 seconds, 2 TN are considered. Additionally, a TN episode is only a consideration if the TN episode is longer than 5 seconds as the average duration of a FoG episode in the collected data was 5.14 seconds (Fig 5). Furthermore, TP, FN, and FP are evaluated from Fig 5 as:

- A single TP is considered if a FoG episode was detected by the algorithm and the patient was clinically labelled as having a FoG episode.
- A single FP is considered if a FoG episode was detected by the algorithm and the patient was clinically labelled as not having a FoG episode.
- A single FN is considered if a FoG episode was not detected by the algorithm and the patient was clinically labelled as having a FoG episode.



**Fig 5. Evaluation method for FoG algorithms.**

doi:10.1371/journal.pone.0171764.g005

### MBFA method assessment

To further analyse the performance of each methodology, a comparison to the most widely used technique of FoG detection is made. Given that the MBFA has been shown to provide good performance (>80% in sensitivity and specificity in a non-user dependent algorithm [27]) and, in addition, the MBFA has been widely used by many authors for detecting FoG [22–24,26,27,31,46–50], the authors consider the MBFA as an accurate baseline method for comparison purposes. Similarly, to our proposed approach, a user-independent (leave-one-patient-out) evaluation has been designed for the MBFA method and compared with the proposed user-independent method described in the *Generic FoG Model Selection* section.

The main parameters to be adjusted within the MBFA method are PTH and FTH. Thus, a range of values for each parameter has been tested for each patient, consisting in values [0,0.5, . . .,4]. After evaluation each pair of PTH and FTH values provide a certain number of TP, TN, FN and FP that are represented in terms of sensitivity ( $Sensitivity = \frac{TP}{TP+FN}$ ) and specificity ( $Specificity = \frac{TN}{TN+FP}$ ). These values are used to evaluate the MBFA algorithm as described in the *Evaluation of the proposed algorithm* section. Thus, let  $Eval_k(PTH, FTH) : \mathbb{R}^2 \rightarrow \mathbb{R}^4$  be a function which returns the number of TP, FP, TN, FN of a given signal from patient  $k$ . From these values, the sensitivity and specificity of the parameters in detecting FoG episodes are obtained based on the previous formulas, which is represented by  $Sens(Eval_k(PTH, FTH))$  and  $Spec(Eval_k(PTH, FTH))$ . Given that a leave-one-patient-out method is employed, a summation for all  $Eval_k(PTH, FTH)$  belonging to all non-testing patients is performed following Eq (5):

$$LOO_t(FTH, PTH) = \sum_{k=1, \dots, 21}^{k \neq t} Eval_k(FTH, PTH) \tag{5}$$

The optimal values for PTH and FTH are obtained according to Eq (6) and by using the data from the training dataset. These values are then evaluated through the testing dataset, which is the data belonging to the testing patient, represented as  $t$  patient in the following

equation:

$$PTH_{opt}, FTH_{opt} = \operatorname{argmax} \left( \sqrt{\operatorname{Sens} (LOO_t(PTH, FTH)) \cdot \operatorname{Spec} (LOO_t(PTH, FTH))} \right) \quad (6)$$

$$s. t. \operatorname{Sens} (LOO_t(PTH, FTH)) > 70\% \wedge \operatorname{Spec}(LOO_t(PTH, FTH)) > 70\%$$

Additionally, a personalised optimisation of the MBFA method has been also performed and compared to the corresponding methodology presented in the *Personalised FoG model selection* section. The user dependent model has also been obtained by including 50% of the episodes from the patient being tested and, in addition, the corresponding windows without episodes (as shown in Fig 5) within the optimisation of PTH and FTH.

Similar to the SVM-based approach presented in the *Personalised FoG model selection* section, the training dataset, which belongs to the patient being tested, is weighted in order to adjust the parameters to the patient’s FoG episodes. Hence, let  $Eval_{t1}(PTH, FTH)$  and  $Eval_{t2}(PTH, FTH)$  be the functions which return the number of TP, FP, TN, and FN from, on the one hand, the first half of the signal containing the first 50% of FoG episodes in the  $t$  patient and, on the other hand, the second half of the signal containing the last 50% of FoG episodes in the  $t$  patient. The evaluation of the remaining part of the signal from the  $t$  patient provides the personalised model performance based on TP, FP, TN and FN, which is represented by  $PM_t(FTH, PTH)$ . The corresponding specificity and sensitivity are obtained according to:

$$\operatorname{Sens}(PM_{t1}(FTH, PTH)) = \operatorname{Sens} (LOO_t(PTH, FTH)) \cdot w_t + \operatorname{Sens}(Eval_{t1}(FTH, PTH)) \cdot w_{20} \quad (7)$$

$$\operatorname{Spec}(PM_{t1}(FTH, PTH)) = \operatorname{Spec} (LOO_t(PTH, FTH)) \cdot w_t + \operatorname{Spec}(Eval_{t1}(FTH, PTH)) \cdot w_{20} \quad (8)$$

Finally, optimal values of PTH and FTH for patient  $t$  in the personalised model are obtained by:

$$PTH_{opt}, FTH_{opt} = \operatorname{argmax} \left( \sqrt{\operatorname{Sens} (PM_{tb}(PTH, FTH)) \cdot \operatorname{Spec} (PM_{tb}(PTH, FTH))} \right) \quad (9)$$

$$s. t. \operatorname{Sens} (PM_{tb}) > 70\% \wedge \operatorname{Spec}(PM_{tb}) > 70\%$$

The same methodology is then applied for  $t2$  in order to evaluate the remaining half of the signals.

## Results and discussion

In this section, the results obtained for the methodologies described in the *Methods* section are reported. To put the results into perspective, the SVM and the MBFA results are shown together. Both SVM and the MBFA methods of FoG detection have been evaluated on a generic model and a personalised model basis. Table 4 presents the specificity, sensitivity and their geometric mean averaged among the 21 PD patients for both methods and both models.

**Table 4. Results for the generic and the personalised models for the proposed SVM-based approach and the MBFA method.** Average specificity, sensitivity and geometric mean, computed as  $\sqrt{\operatorname{Specificity} \cdot \operatorname{Sensitivity}}$ , for the 21 PD patients are presented.

	Generic model			Personalised model		
	Sensitivity	Specificity	Geometric mean	Sensitivity	Specificity	Geometric mean
SVM	74.7%	79.0%	76.8%	88.1%	80.1%	84.0%
MBFA	81.6%	52.6%	65.6%	89.1%	61.5%	74.0%

doi:10.1371/journal.pone.0171764.t004

The values reported in [Table 4](#) have been obtained by summing the TP, TN, FP and FN values and solving for sensitivity and specificity as reported in the previous sections.

[Table 4](#) clearly shows that the personalised models for both the SVM and the MBFA methods outperform the generic models, providing an increase of 7.2% for SVM and 8.4% for MBFA in terms of geometric mean. Thus, the detection of FoG episodes are significantly improved when data from the patient being tested is used to train the model or to adjust the MBFA parameters. The resulting increase of 8.4% obtained for the MBFA is consistent with the results reported in [21]. Similarly, the SVM method shows better results by considering the data from the patient being tested, which is an interesting finding that has not been previously reported and that can help to obtain enhanced models for FoG detection.

Regarding the performance of the MBFA and the SVM method, it can be seen that the SVM methods geometric mean is significantly greater than the MBFA methods geometric mean. Specifically, the proposed method outperforms the MBFA method with an increase of 11.2% for the generic model and an increase of 10% for the personalised model. Thus, the detection of FoG episodes are significantly improved using the SVM method in comparison to the MBFA method. In fact, the machine learning approach better generalises the problem in all the results obtained. This is due to the inclusion of several features that characterise in a more accurate way not only the FoG episodes but also those parts with possible FP, thus providing a much higher specificity. In this sense, it is notable that the specificity of the MBFA generic model is 52.6%, while the corresponding SVM-based specificity is 79%.

[Table 5](#) presents the individual patient results obtained by the SVM method for both the generic and the personalised models. These results aim to show how the personalisation model improves the detection of FoG episodes in comparison to the generic model. For example, the geometric mean for patient 1 is only 37.35% for the generic model and 86.2% for the personalisation model. In total, for 16 patients the personalisation provides an increase in the geometric mean which corresponds to a greater detection of FoG episodes. However, in the remaining 5 patients (number 4, 10, 11, 12 and 18), the geometric mean is slightly lower, below 3%. This may occur when the scenario in the first half of the test is different to the second part and, thus, episodes may appear differently. Most patients freeze similarly during the whole test but, however, there are patients who freeze in a different manner inside their homes than outside. Trained models can learn that, in the main, episodes occur in a stereotypical way. When the evaluation is performed outdoors or in a different scenario in the other half of the test the classifier might not work properly and thus the mean could decrease a little compared to the generic SVM.

Some previous papers have reported results of sensitivity and specificity over 90%, which are higher than those shown in this paper. However, some issues are evident in those works [28,48,52].

The evaluation method consists of a leave one window out, which actually trains the model with all the windows (including the testing patient) except for one. Therefore, the testing signals are very similar to the trained signals, resulting in very high accuracies. In addition, the leaving one window out approach considers each window either as a positive or a negative case, which overestimates specificity and sensitivity values, as presented in the previous sections. In the majority of the papers, results have been obtained by considering either laboratory conditions or the use of multiple sensors, which increases the accuracy of FoG detection.

In this work, a more reliable evaluation method has been employed, which consists in evaluating through the leave one patient out method. Using this method, at least half of the signals from the testing patient (for the personalised model) has been omitted in order to properly test the model. Additionally, our results have been obtained in the patients' own home and by using a single sensor located on the waist. The waist location has been shown to be comfortable

**Table 5. Results of the proposed SVM-based approach for both generic and personalised models.**

Index	Generic model			Personalised model		
	Sensitivity	Specificity	Geometric mean	Sensitivity	Specificity	Geometric mean
Patient 1	18.89%	73.85%	37.35%	92.08%	80.70%	86.20%
Patient 2	46.77%	51.88%	49.26%	80.95%	59.81%	69.58%
Patient 3	39.58%	92.21%	60.41%	100.00%	90.91%	95.35%
Patient 4	100.00%	60.00%	77.46%	96.55%	57.89%	74.77%
Patient 5	92.59%	83.05%	87.69%	98.28%	89.29%	93.67%
Patient 6	100.00%	68.99%	83.06%	97.50%	83.04%	89.98%
Patient 7	86.96%	86.36%	86.66%	91.67%	83.33%	87.40%
Patient 8	100.00%	57.01%	75.50%	96.72%	73.91%	84.55%
Patient 9	83.33%	53.73%	66.91%	80.00%	75.00%	77.46%
Patient 10	93.59%	73.49%	82.94%	84.88%	77.01%	80.85%
Patient 11	93.31%	92.59%	92.95%	94.67%	88.10%	91.32%
Patient 12	83.95%	60.98%	71.55%	77.91%	61.04%	68.96%
Patient 13	91.25%	66.41%	77.84%	86.42%	77.34%	81.76%
Patient 14	88.14%	64.91%	75.64%	81.97%	74.77%	78.29%
Patient 15	98.73%	77.78%	87.63%	96.20%	82.35%	89.01%
Patient 16	92.73%	62.34%	76.03%	91.07%	71.43%	80.65%
Patient 17	70.31%	91.67%	80.28%	78.79%	91.49%	84.90%
Patient 18	100.00%	88.89%	94.28%	92.31%	92.21%	92.26%
Patient 19	82.35%	93.85%	87.91%	92.31%	96.88%	94.56%
Patient 20	32.35%	75.47%	49.41%	51.43%	85.71%	66.39%
Patient 21	64.71%	92.68%	77.44%	88.24%	89.74%	88.99%
<b>Average</b>	<b>79.03%</b>	<b>74.67%</b>	<b>75.15%</b>	<b>88.09%</b>	<b>80.09%</b>	<b>83.66%</b>

doi:10.1371/journal.pone.0171764.t005

for users [32] but, in contrast, the detection of FoG accuracy is not as accurate as that obtained by setting the inertial systems on the feet [21]. In consequence, the additional difficulty provided through these conditions also explain the lower results obtained.

With a specificity of 61% and sensitivity values of approximately 50% on average, it is evident that the evaluation of MBFA is reduced due to the new evaluation methodology. In addition, the effect of performing the dataset in patient’s home environments and while performing some false positive tests, suggests the MBFA may not be very reliable in a home environment.

The proposed evaluation method assesses the algorithm for FoG detection in relation to the number of episodes, by balancing the number of TN, TP, FN and FP. However, this method is not able to determine the length of each episode. Given that the processing method is performed through windows, the resolution of the episode can only be 1.6 seconds and the minimum FoG episode time is 3.2 seconds according to the specifications reported in the *Pre-Processing and feature extraction* section.

So far, the MBFA model has been seen as an optimal model due to its low computational cost and good performance. Compared to our model, the computational burden is significantly lower; nonetheless, the accuracies obtained are also lower. However, with recent technological advances, the inclusion of virtual digital signal processors within the microcontrollers and the relevant growth of performance with low power consumption makes it feasible to implement real-time machine learning classifiers with several support vectors and wide feature vectors. In spite of this, the authors considered that the models obtained by the proposed approach should be optimised in the future before being implemented in real time. The model

purposed has been optimised by maximising the product of sensitivity and specificity but not including any constraint to minimise support vectors and, in addition, it has no attempt to select features among the 55 proposed features undertaken. Regarding the online optimisation, the non-user dependent model presented employs an average of  $6790 \pm 387$  support vectors among the different leave-one-out models. Considering 4 bytes for each floating point number, an average of 1494KB of memory is required in order to store the SVM model. On the other hand, the user dependent model uses  $9221 \pm 520$  support vectors requiring 2028KB. These models may require a considerable portion of the memory available on some low to moderate power consumption and high-performance microcontroller such as Cortex-M4 ARM microcontrollers. Thus, a reduction of support vectors may be done, for instance, by means of Separable Case Approximation method [73]. Using this method, the non-user dependent model is reduced from 7147 support vectors to 4598, reducing the amount of memory needed by 30%. Although the proposed algorithm has been developed to be implemented online within a microcontroller, and there are low power high-performance microcontrollers that could support the required 1011KB (corresponding memory for 4598 support vectors), a reduction of features should be done, together with further removal of support vectors, in order to minimise the computational burden on the prediction phase. However, this may reduce performance of the algorithm, thus, the trade-off between performance and computational burden, which directly affects the microcontroller capacity and increases its consumption has to be considered.

## Conclusions

The MBFA is a method to detect FoG episodes with the analysis of frequency features obtained by an accelerometer signal of a wearable system. This method has been used in numerous studies [22–24,26,27,31,46–50] due to its simplicity (just the optimisation of 2 features) and its high accuracy in detecting FoG episodes (>80%, according to Moore et al. [27]). However, the method has only been tested under laboratory conditions without taking into account situations which could lead to false positives. As far as the authors know, the evaluation method does not consider those situations where long periods of negative outputs such as sitting, lying, or standing, overstate the specificity results.

This paper presents a supervised machine learning approach to detect FoG that has been trained and evaluated with an inertial dataset obtained from an IMU located at patients' waist while the PD patients were at home [37]. Two methodologies have been proposed and evaluated based on the leave one out method. On the one hand, a generic model that is not adjusted to any patient has been proposed, and, on the other hand, a personalised model, which is trained with part of the inertial signal from the patient undergoing testing. Both generic and personalised model selection approaches have been assessed with the MBFA and the proposed machine learning method. Furthermore, an optimisation method to select the MBFA model has been employed. The evaluation has been performed by balancing the number of true positives and true negatives and also evaluating episode instead of the window, leading to a more reliable assessment. For the personalised model, only the 50% of the FoG episodes including those non-episodic signals have been used for training, enabling the rest of the signals to be used for evaluation. This percentage has been selected given that some patients have a low number of FoG episodes, this way, the evaluation set is larger. Furthermore, validation of results is more significant if evaluation percentage is larger since the classifier avoids overfitting, reinforcing, thus, the performance of the model.

Our results show how the personalised models outperform the generic models by on average by 10%. However, and as stated in the *Personalised FoG model selection* section, it requires

a data capture and data training process, which is not necessary with the non-user dependent model. This process involves taking data from patients with video recording to capture different FoG episodes at the patients' home with a specific protocol test to elicit FoG episodes. This method of data collection can be cumbersome to obtain for the patient.

On the other hand, the machine learning approach enhances significantly the results obtained with the MBFA model. Evaluations per episode show that the machine learning approach either increases both sensitivity and specificity or may only increase one, leave the other unchanged. The geometric mean of both sensitivity and specificity and their average have been used to compare both models (MBFA and SVM). SVM enhances these values by 8% and 11%, respectively. Improvements might be explained by 2 factors. Firstly, the capacity of generalisation of a machine learning method in comparison to the 2 thresholds based MBFA method. Secondly, is the inclusion of several features, which better characterise the episodes and the non-episodic movements. Since the database was constructed from data collected in real environments, situations that are more reflective of real conditions could be used in order to strengthen the classifier against false positives, achieving much better specificities. However, the algorithm needs some optimisation in order to be implemented in a microcontroller of a wearable device for real-time applications. The most relevant issue is the number of support vectors, which should be reduced as shown in the *Evaluation of the proposed algorithm* section. This reduction would lower the computational burden.

Results obtained provide a sensitivity and specificity geometric mean of 84% in the personalised model and of 76% in the generic model, which are significant enough to evaluate FoG at patients' home. These results have been obtained under challenging conditions, at patients' home and with a rigorous evaluation process in order to balance the number of true positives and true negatives.

As far as authors know, this algorithm is the first to achieve high sensitivities and specificities considering an online implementation using a balanced assessment per episode through patients' home training, which included a false positive test. This algorithm opens up the possibility to rigorously assess FoG in PD patients in real time with the aim of acting quickly with some cueing modalities [8,74]. Additionally, given that medication considerably affects FoG frequency and severity in some patients [63], this algorithm could help to accurately tailor medication regimes in future.

## Acknowledgments

Part of this project has been performed within the framework of the MASPARK project which is funded by La Fundació La Marató de TV3 20140431. This work also forms part of the framework of the FP7 project REMPARK ICT-287677, which is funded by the European Community. The authors, thus, would like to acknowledge the contributions of their colleagues from MASPARK and REMPARK consortium.

## Author Contributions

**Conceptualization:** DRM AS CPL A. Catala AB TJC HL RA ARM.

**Data curation:** SA BM AP MCC AB.

**Formal analysis:** DRM AS CPL.

**Funding acquisition:** JC JMM HL AB GO.

**Investigation:** AB SA BM TJC PB LRQ DS GV RA A. Costa ARM.



**Methodology:** ARM RA GV HL JA TJC MCC AP AB JMM CPL AS DRM.

**Project administration:** JC JMM HL JA AB GO.

**Resources:** DRM AS CPL A. Catala JMM JC AB TJC GV RA.

**Software:** CPL DRM AS.

**Supervision:** JC JMM ARM AB.

**Validation:** AS CPL DRM AB.

**Visualization:** DRM AS CPL.

**Writing – original draft:** DRM AS CPL.

**Writing – review & editing:** AS CPL DRM A. Catala JMM JC TJC PB LRQ GO A. Costa ARM.

## References

1. European Parkinson's Disease Association, Working Group on Parkinson's Disease (World Health Organization). Prevalence of Parkinson's Disease. 2015. <http://www.epda.eu.com/en/resources/life-with-parkinsons/part-1/prevalence-of-parkinsons-disease/>
2. Fahn S. The freezing phenomenon in parkinsonism. *Adv Neurol*. 1995; 67: 53–63. Available: <http://www.ncbi.nlm.nih.gov/pubmed/8848982> PMID: 8848982
3. Giladi N, Treves Ta, Simon ES, Shabtai H, Orlov Y, Kandinov B, et al. Freezing of gait in patients with advanced Parkinson's disease. *J Neural Transm*. Springer; 2001; 108: 53–61.
4. Bloem BR, Hausdorff JM, Visser JE, Giladi N. Falls and freezing of gait in Parkinson's disease: a review of two interconnected, episodic phenomena. *Mov Disord*. 2004; 19: 871–884. doi: [10.1002/mds.20115](https://doi.org/10.1002/mds.20115) PMID: 15300651
5. Moreau C, Defebvre L, Bleuse S, Blatt JL, Duhamel A, Bloem BR, et al. Externally provoked freezing of gait in open runways in advanced Parkinson's disease results from motor and mental collapse. *J Neural Transm*. 2008; 115: 1431–1436. doi: [10.1007/s00702-008-0099-3](https://doi.org/10.1007/s00702-008-0099-3) PMID: 18726136
6. Lieberman A. Are Freezing of Gait (FOG) and panic related? *J Neurol Sci*. 2006; 248: 219–222. doi: [10.1016/j.jns.2006.05.023](https://doi.org/10.1016/j.jns.2006.05.023) PMID: 16797596
7. Aarsland D, Brønnick K, Ehrt U, De Deyn PP, Tekin S, Emre M, et al. Neuropsychiatric symptoms in patients with Parkinson's disease and dementia: frequency, profile and associated care giver stress. *J Neurol Neurosurg Psychiatry*. 2007; 78: 36–42. doi: [10.1136/jnnp.2005.083113](https://doi.org/10.1136/jnnp.2005.083113) PMID: 16820421
8. Lim I, van Wegen E, de Goede C, Deutekom M, Nieuwboer A, Willems A, et al. Effects of external rhythmic cueing on gait in patients with Parkinson's disease: a systematic review. *Clin Rehabil*. 2005; 19: 695–713. PMID: 16250189
9. Bächlin M, Plotnik M, Roggen D, Giladi N, Hausdorff JM, Tröster G. A wearable system to assist walking of Parkinson's disease patients. *Methods Inf Med*. 2010; 49: 88–95. doi: [10.3414/ME09-02-0003](https://doi.org/10.3414/ME09-02-0003) PMID: 20011807
10. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc*. 1991; 39: 142–8. PMID: 1991946
11. Schaafsma JD, Balash Y, Gurevich T, Bartels AL, Hausdorff JM, Giladi N. Characterization of freezing of gait subtypes and the response of each to levodopa in Parkinson's disease. *Eur J Neurol*. 2003; 10: 391–398. PMID: 12823491
12. Snijders AH, Haaxma CA, Hagen YJ, Munneke M, Bloem BR. Freezer or non-freezer: clinical assessment of freezing of gait. *Parkinsonism Relat Disord*. 2012; 18: 149–54. doi: [10.1016/j.parkreldis.2011.09.006](https://doi.org/10.1016/j.parkreldis.2011.09.006) PMID: 21968033
13. Nonnekes J, Janssen AM, Mensink SHG, Oude Nijhuis LB, Bloem BR, Snijders AH. Short rapid steps to provoke freezing of gait in Parkinson's disease. *J Neurol*. 2014; 261: 1763–1767. doi: [10.1007/s00415-014-7422-8](https://doi.org/10.1007/s00415-014-7422-8) PMID: 24957299
14. Giladi N, Shabtai H, Simon E, Biran S, Tal J, Korczyn A. Construction of freezing of gait questionnaire for patients with Parkinsonism. *Parkinsonism Relat Disord*. 2000; 6: 165–170. PMID: 10817956

15. Giladi N, Tal J, Azulay T, Rascol O, Brooks DJ, Melamed E, et al. Validation of the freezing of gait questionnaire in patients with Parkinson's disease. *Mov Disord*. 2009; 24: 655–661. doi: [10.1002/mds.21745](https://doi.org/10.1002/mds.21745) PMID: [19127595](https://pubmed.ncbi.nlm.nih.gov/19127595/)
16. Nilsson MH, Hariz G-M, Wictorin K, Miller M, Forsgren L, Hagell P. Development and testing of a self administered version of the Freezing of Gait Questionnaire. *BMC Neurol*. 2010; 10: 85. doi: [10.1186/1471-2377-10-85](https://doi.org/10.1186/1471-2377-10-85) PMID: [20863392](https://pubmed.ncbi.nlm.nih.gov/20863392/)
17. Nieuwboer A, De Weerd W, Dom R, Lesaffre E. A frequency and correlation analysis of motor deficits in Parkinson patients. *Disabil Rehabil*. 1998; 20: 142–150. PMID: [9571381](https://pubmed.ncbi.nlm.nih.gov/9571381/)
18. Papapetropoulos SS. Patient Diaries As a Clinical Endpoint in Parkinson's Disease Clinical Trials. *CNS Neurosci Ther*. 2012; 18: 380–387. doi: [10.1111/j.1755-5949.2011.00253.x](https://doi.org/10.1111/j.1755-5949.2011.00253.x) PMID: [22070400](https://pubmed.ncbi.nlm.nih.gov/22070400/)
19. Han JH, Lee WJ, Tae Beom A, Beom Suk J, Kwang Suk P. Gait analysis for freezing detection in patients with movement disorder using three dimensional acceleration system. 25th Annual International Conference of the IEEE EMBS. 2003. pp. 1863–1865.
20. Moore ST, MacDougall HG, Ondo WG. Ambulatory monitoring of freezing of gait in Parkinson's disease. *J Neurosci Methods*. 2008; 167: 340–348. doi: [10.1016/j.jneumeth.2007.08.023](https://doi.org/10.1016/j.jneumeth.2007.08.023) PMID: [17928063](https://pubmed.ncbi.nlm.nih.gov/17928063/)
21. Bächlin M, Roggen D, Plotnik M, Hausdorff JM, Giladi N, Tröster G. Online Detection of Freezing of Gait in Parkinson's Disease Patients: A Performance Characterization. *Proc 4th Int ICST Conf Body Area Networks*. Icast; 2009;
22. Jovanov E, Wang E, Verhagen L, Fredrickson M, Fratangelo R. deFOG—A real time system for detection and unfreezing of gait of Parkinson's patients. 31st Annual International Conference of the IEEE EMBS. 2009. pp. 5151–5154.
23. Delval A, Snijders AH, Weerdesteijn V, Duysens JE, Defebvre L, Giladi N, et al. Objective detection of subtle freezing of gait episodes in Parkinson's disease. *Mov Disord*. 2010; 25: 1684–1693. doi: [10.1002/mds.23159](https://doi.org/10.1002/mds.23159) PMID: [20544715](https://pubmed.ncbi.nlm.nih.gov/20544715/)
24. Zabaleta H, Keller T, Fimbel EJ. Gait analysis in frequency domain for freezing detection in patients with Parkinson's disease. *Gerontechnology*. 2008; 7: 4.
25. Cole BT, Roy SH, Nawab SH. Detecting Freezing-of-Gait During Unscripted and Unconstrained Activity. 33rd Annual International Conference of the IEEE EMBS. 2011. pp. 5649–5652.
26. Niazmand K, Tonn K, Zhao Y, Fietzek UM, Schroeteler F, Ziegler K, et al. Freezing of Gait detection in Parkinson's disease using accelerometer based smart clothes. *IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE; 2011. pp. 201–204.
27. Moore ST, Yungheer Da, Morris TR, Dilda V, MacDougall HG, Shine JM, et al. Autonomous identification of freezing of gait in Parkinson's disease from lower-body segmental accelerometry. *J Neuroeng Rehabil*. *Journal of NeuroEngineering and Rehabilitation*; 2013; 10: 19. doi: [10.1186/1743-0003-10-19](https://doi.org/10.1186/1743-0003-10-19) PMID: [23405951](https://pubmed.ncbi.nlm.nih.gov/23405951/)
28. Ahlrichs C, Samà A, Lawo M, Cabestany J, Rodríguez-Martín D, Pérez-López C, et al. Detecting freezing of gait with a tri-axial accelerometer in Parkinson's disease patients. *Med Biol Eng Comput*. 2016; 54: 223–233. doi: [10.1007/s11517-015-1395-3](https://doi.org/10.1007/s11517-015-1395-3) PMID: [26429349](https://pubmed.ncbi.nlm.nih.gov/26429349/)
29. Bächlin M, Plotnik M, Roggen D, Maidan I, Hausdorff JM, Giladi N, et al. Wearable assistant for Parkinsons disease patients with the freezing of gait symptom. *IEEE Trans Inf Technol Biomed*. 2010; 14: 436–446. doi: [10.1109/TITB.2009.2036165](https://doi.org/10.1109/TITB.2009.2036165) PMID: [19906597](https://pubmed.ncbi.nlm.nih.gov/19906597/)
30. Weiss A, Herman T, Giladi N, Hausdorff JM. New evidence for gait abnormalities among Parkinson's disease patients who suffer from freezing of gait: insights using a body-fixed sensor worn for 3 days. *J Neural Transm*. 2015; 122: 403–410. doi: [10.1007/s00702-014-1279-y](https://doi.org/10.1007/s00702-014-1279-y) PMID: [25069586](https://pubmed.ncbi.nlm.nih.gov/25069586/)
31. Zach H, Janssen AM, Snijders AH, Delval A, Ferraye MU, Auff E, et al. Identifying freezing of gait in Parkinson's disease during freezing provoking tasks using waist-mounted accelerometry. *Parkinsonism Relat Disord*. Elsevier Ltd; 2015; 21: 1362–1366.
32. Mathie MJ, Basilakis J, Celler BG. A system for monitoring posture and physical activity using accelerometers. 23rd Annu Int Conf IEEE EMBS. IEEE; 2001;4: 3654–3657.
33. Yang CC, Hsu YL. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*. 2010; 10: 7772–7788. doi: [10.3390/s100807772](https://doi.org/10.3390/s100807772) PMID: [22163626](https://pubmed.ncbi.nlm.nih.gov/22163626/)
34. Pérez-López C, Samà A, Rodríguez-Martín D, Català A, Cabestany J, de Mingo E, et al. Monitoring Motor Fluctuations in Parkinson's Disease Using a Waist-Worn Inertial Sensor. *International Work Conference on Artificial Neural Networks Advances in Computational Intelligence Lecture Notes on Computer Science*. 2015. pp. 461–474.
35. Pérez-López C, Samà A, Rodríguez-Martín D, Moreno-Aróstegui JM, Cabestany J, Bayes A, et al. Dopaminergic-induced dyskinesia assessment based on a single belt-worn accelerometer. *Artif Intell Med*. Elsevier B.V.; 2016; 67: 47–56. doi: [10.1016/j.artmed.2016.01.001](https://doi.org/10.1016/j.artmed.2016.01.001) PMID: [26831150](https://pubmed.ncbi.nlm.nih.gov/26831150/)

36. MASPARK. Freezing in Parkinson's Disease: Improving Quality of Life with an Automatic Control System- La Marató de TV3-20140431. 2015–2017.
37. Samà A, Perez-Lopez C, Rodriguez-Martin D, Cabestany J, Moreno JM, Rodriguez-Molinero A. A Heterogeneous Database for Movement Knowledge Extraction in Parkinson's Disease. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. 2013.
38. Hausdorff JM, Balash Y, Giladi N. Time series analysis of leg movements during freezing of gait in Parkinson's disease: akinesia, rhyme or reason? *Physica A*. 2003; 321: 565–570.
39. Nieuwboer A, De Weerdts W, Desloovere K, Fieuw S, Broens-kaucsik E. Abnormalities of the Spatio-temporal Characteristics of Gait at the Onset of Freezing in Parkinson's Disease. *Mov Disord*. 2001; 16: 1066–1075. PMID: [11748737](#)
40. Knobl P, Kielstra L, Almeida Q. The relationship between motor planning and freezing of gait in Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2012; 83: 98–101. doi: [10.1136/jnnp-2011-300869](#) PMID: [21836031](#)
41. Almeida QJ, Lebold CA. Freezing of gait in Parkinson's disease: a perceptual cause for a motor impairment? *J Neurol Neurosurg Psychiatry*. 2010; 81: 513–518. doi: [10.1136/jnnp.2008.160580](#) PMID: [19758982](#)
42. Handojoseno a MA, Shine JM, Nguyen TN, Tran Y, Lewis SJG, Nguyen HT. The detection of Freezing of Gait in Parkinson's disease patients using EEG signals based on Wavelet decomposition. 34th Annu Int Conf IEEE Eng Med Biol Soc. 2012;2012: 69–72.
43. Mazilu S, Calatroni A, Gazit E, Mirelman A, Hausdorff J, Troester G. Prediction of Freezing of Gait in Parkinson's from Physiological Wearables: An Exploratory Study. *IEEE J Biomed Heal Informatics*. 2015; 6: 1–1.
44. Takač B, Català A, Rodríguez Martín D, van der Aa N, Chen W, Rauterberg M. Position and Orientation Tracking in a Ubiquitous Monitoring System for Parkinson Disease Patients With Freezing of Gait Symptom. *JMIR mhealth uhealth*. 2013; 1: e14. doi: [10.2196/mhealth.2539](#) PMID: [25098265](#)
45. Weiss A, Herman T, Giladi N, Hausdorff JM. Objective assessment of fall risk in Parkinson's disease using a body-fixed sensor worn for 3 days. *PLoS One*. 2014; 9: e96675. doi: [10.1371/journal.pone.0096675](#) PMID: [24801889](#)
46. Zhao Y, Tonn K, Niazmand K, Fietzek UM, D'Angelo LT, Ceballos-Baumann a., et al. Online FOG Identification in Parkinson's disease with a time-frequency combined Algorithm. *Proc 2012 IEEE-EMBS Int Conf Biomed Heal Informatics*. 2012;25: 192–195.
47. Mancini M, Priest KC, Nutt JG, Horak FB. Quantifying freezing of gait in Parkinson's disease during the instrumented timed up and go test. 34th Annu Int Conf IEEE Eng Med Biol Soc. 2012;2012: 1198–201.
48. Mazilu S, Hardegger M, Zhu Z, Roggen D, Tröster G, Plotnik M, et al. Online Detection of Freezing of Gait with Smartphones and Machine Learning Techniques. *Proc 6th Int ICST Conf Pervasive Comput Technol Healthc*. 2012; 123–130.
49. Mazilu S, Calatroni A, Gazit E, Roggen D, Hausdorff JM, Tröster G. Feature Learning for Detection and Prediction of Freezing of Gait in Parkinson's Disease. *Proceedings of the 9th international conference on Machine Learning and Data Mining in Pattern Recognition*. 2013. pp. 144–158.
50. Rodríguez-Martín D, Samà A, Pérez-López C, Català A, Cabestany J, Rodríguez-Molinero A. Posture Detection with waist-worn Accelerometer: An application to improve Freezing of Gait detection in Parkinson's disease patients. *European Conference on Ambient Assisted Living*. Eindhoven; 2014. pp. 1–6.
51. Stamatakis J, Crémers J, Maquet D, Macq B, Garraux G. Gait feature extraction in Parkinson's disease using low-cost accelerometers. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society*. 2011. pp. 7900–7903.
52. Tripoliti EE, Tzallas AT, Tsipouras MG, Rigas G, Bougia P, Leontiou M, et al. Automatic detection of freezing of gait events in patients with Parkinson's disease. *Comput Methods Programs Biomed*. Elsevier Ireland Ltd; 2013; 110: 12–26.
53. Coste C, Sijobert B, Pissard-Gibollet R, Pasquier M, Espiau B, Geny C. Detection of Freezing of Gait in Parkinson Disease: Preliminary Results. *Sensors*. 2014; 14: 6819–6827. doi: [10.3390/s140406819](#) PMID: [24740014](#)
54. Tay A, Yen SC, Lee PY, Wang CY, Neo A, Phan SW, et al. Freezing of Gait (FoG) detection for Parkinson Disease. 10th Asian Control Conference (ASCC). 2015. pp. 1–6.
55. REMPARK—Personal Health Device for the Remote and Autonomous Management of Parkinson's Disease. FP7-ICT-2011-7-287677. 2011–2014.
56. Bachlin M. DAPHNET Project-Dynamic Analysis of Physiological Networks. Sixth Framework Programme, FP6. <http://www.wearable.ethz.ch/research/groups/health/daphnet/index>

57. Maidan I, Plotnik M, Mirelman A, Weiss A, Giladi N, Hausdorff JM. Heart rate changes during freezing of gait in patients with Parkinson's disease. *Mov Disord*. 2010; 25: 2346–2354. doi: [10.1002/mds.23280](https://doi.org/10.1002/mds.23280) PMID: [20721914](https://pubmed.ncbi.nlm.nih.gov/20721914/)
58. Takač B, Català A, Rodríguez Martín D, van der Aa N, Chen W, Rauterberg M. Position and Orientation Tracking in a Ubiquitous Monitoring System for Parkinson Disease Patients With Freezing of Gait Symptom. *JMIR mhealth uhealth*. *JMIR mhealth and uhealth*; 2013; 1: e14. doi: [10.2196/mhealth.2539](https://doi.org/10.2196/mhealth.2539) PMID: [25098265](https://pubmed.ncbi.nlm.nih.gov/25098265/)
59. Maidan I, Bernad-Elazari H, Gazit E, Giladi N, Hausdorff JM, Mirelman A. Changes in oxygenated hemoglobin link freezing of gait to frontal activation in patients with Parkinson disease: an fNIRS study of transient motor-cognitive failures. *J Neurol*. 2015; 262: 899–908. doi: [10.1007/s00415-015-7650-6](https://doi.org/10.1007/s00415-015-7650-6) PMID: [25636682](https://pubmed.ncbi.nlm.nih.gov/25636682/)
60. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinicopathological study of 100 cases. *J Neurol Neurosurg Psychiatry*. 1992; 55: 181–184. PMID: [1564476](https://pubmed.ncbi.nlm.nih.gov/1564476/)
61. Cabestany J, Moreno JM, Perez C, Sama A, Catala A. REMPARK: When AI and Technology Meet Parkinson Disease Assessment. 20th International Conference on Mixed Design of Integrated Circuits and Systems. 2013.
62. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology*. 1967; 17: 427–427. PMID: [6067254](https://pubmed.ncbi.nlm.nih.gov/6067254/)
63. Amboni M, Barone P, Picillo M, Cozzolino A, Longo K, Erro R, et al. A two-year follow-up study of executive dysfunctions in Parkinsonian patients with freezing of gait at on-state. *Mov Disord*. 2010; 25: 800–802. doi: [10.1002/mds.23033](https://doi.org/10.1002/mds.23033) PMID: [20198648](https://pubmed.ncbi.nlm.nih.gov/20198648/)
64. Rodríguez-Martín D, Pérez-López C, Samà A, Cabestany J, Català A. A Wearable Inertial Measurement Unit for Long-Term Monitoring in the Dependency Care Area. *Sensors*. 2013; 13: 14079–14104. doi: [10.3390/s131014079](https://doi.org/10.3390/s131014079) PMID: [24145917](https://pubmed.ncbi.nlm.nih.gov/24145917/)
65. Zhou H, Hu H. Human motion tracking for rehabilitation—A survey. *Biomed Signal Process Control*. 2008; 3: 1–18.
66. Rodríguez-Martín D, Samà A, Pérez-López C, Català A, Cabestany J, Rodríguez-Molinero A. SVM-based posture identification with a single waist-located triaxial accelerometer. *Expert Syst Appl*. Elsevier Ltd; 2013; 40: 7203–7211.
67. Najafi B, Aminian K, Paraschiv-Ionescu A, Loew F, Bula CJ, Robert P. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *IEEE Trans Biomed Eng*. 2003; 50: 711–723. doi: [10.1109/TBME.2003.812189](https://doi.org/10.1109/TBME.2003.812189) PMID: [12814238](https://pubmed.ncbi.nlm.nih.gov/12814238/)
68. Rodríguez-Martín D, Samà A, Pérez-López C, Cabestany J, Català A, Rodríguez-Molinero A. Posture transition identification on PD patients through a SVM-based technique and a single waist-worn accelerometer. *Neurocomputing*. Elsevier; 2015; 164: 144–153.
69. Khan A M, Lee YK, Kim TS. Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets. *Conf Proc IEEE Eng Med Biol Soc*. 2008;2008: 5172–5175.
70. Boni A, Pianegiani F, Petri D. Low-Power and Low-Cost Implementation of SVMs for Smart Sensors. *IEEE Trans Instrum Meas*. 2007; 56: 39–44.
71. Vapnik VN. *The Nature of Statistical Learning Theory*. second. New York: Springer-Verlag; 1995.
72. Rodríguez-Martín D, Samà A, Pérez-López C, Català A, Cabestany J, Rodríguez-Molinero A. Posture Detection with waist-worn Accelerometer: An application to improve Freezing of Gait detection in Parkinson's disease patients. *Recent Advances in Ambient Assisted Living—Bridging Assistive Technologies, e-Health and Personalized Health Care*. 1st ed. 2015. pp. 3–17.
73. Geebelen D, Suykens JAK, Vandewalle J. Reducing the number of support vectors of SVM classifiers using the smoothed separable case approximation. *IEEE Trans Neural Networks Learn Syst*. 2012; 23: 682–688.
74. Arias P, Cudeiro J. Effect of rhythmic auditory stimulation on gait in Parkinsonian patients with and without freezing of gait. *PLoS One*. 2010; 5: 8.