

# Fall Risk Prediction via Classification of Lower Extremity Strength in Older Adults with Exergame-Collected Data

Hagen Becker<sup>1</sup>, Augusto Garcia-Agundez<sup>1</sup>, Philipp Niklas Müller<sup>1</sup>, Thomas Tregel<sup>1</sup>, André Miede<sup>2</sup> and Stefan Göbel<sup>1</sup>

<sup>1</sup> Multimedia Communications Lab, Technische Universität Darmstadt, Darmstadt, Germany

<sup>2</sup> HTW Saar, Saarbrücken University of Applied Sciences, Saarbrücken, Germany

Corresponding author: Augusto Garcia-Agundez, Multimedia Communications Lab, TU Darmstadt, Rundeturmstr. 10, 64283 Darmstadt, Germany. Augusto.garcia@kom.tu-darmstadt.de

**Abstract. Objective:** The goal of this article is to present and evaluate a sensor-based falling risk estimation system. The system consists of an array of Wii Balance Boards (WBB) and an exergame that estimates if the player is at an increased falling risk by predicting the result of the 30 Second Sit-To-Stand Test (30CST). **Methods:** 16 participants recruited at a nursing home performed the 30CST and then played the exergame as often as desired during a period of two weeks. For each session, features related to how they walk and stand on the WBBs while playing the exergame were collected. Different classifier algorithms were used to predict the result of the 30CST on a binary basis (able or unable to maintain physical independence). **Results:** We achieved a maximum accuracy of 91% when attempting to estimate if the player's 30CST score will be over or under a threshold of 12 points using a Logistic Model Tree. We also believe it is feasible to predict age- and sex-adjusted cutoff scores. **Conclusion:** An array of WBBs seems to be a viable solution to estimate lower extremity strength and with it the falling risk. In addition, data extracted while playing may form a basis to perform a general screening to identify elderly at an increased falling risk.

**Keywords.** Wii Balance Board, Fall Risk, Fall Detection, Balance, Exergames, Serious Games

1 **Introduction**

2

3 Among older adults, falls are an important cause of mortality and early placement in nursing homes. The  
4 main causes of falls are accidental and environment-related (31%) or caused by gait imbalances (17%).  
5 Approximately 30 to 60% of older adults fall each year. From these falls, 10 to 20% result in injury,  
6 hospitalization, or death. Among the most relevant factors to prevent these falls are risk assessment and  
7 exercise [2].

8

9 In the field of training and health, exergames, active video games that incorporate physical movements,  
10 aim to combine physical exercise with the fun associated with gaming. The main advantage of this approach  
11 is the possibility of designing games that perform fall risk prevention exercises or collect clinically  
12 meaningful data in the background, while the user plays the exergame. An additional advantage of this  
13 approach, in comparison to traditional exercise, is the possibility to adapt the exergame to the specific  
14 needs of the user in real time and without external intervention based on game data [3].

15

16 The possibility of using the Wii Balance Board (WBB) to estimate whether the player is at an increased  
17 falling risk has been identified [4]. However, related publications mention the need for additional studies,  
18 particularly in finding direct relationships between sensor data and clinically meaningful fall risk estimation  
19 methods. Studies show that WBB data contain information that allow a discrimination between elderly who  
20 previously fell and others who did not [5]. Early evidence also shows that the WBB could be used to train  
21 balance in the elderly [6], and that there are statistically significant differences in the way elderly at falling  
22 risk interact with the WBB. These differences actually correlate with clinical fall risk tests [7]. Table 1  
23 summarizes the specific details of these studies. A limitation of the WBB is that, due to its small surface, it  
24 can only be used to estimate balance while standing, and not in movement.

25

26

27

28

29

Table 1: Recent Studies using WBB data to identify players at falling risk

Author	N	Data	Main goal	Results
Mertes et al. [5]	12	Center of pressure	Classify elderly who fell in the past and others who did not	76.6% Classification accuracy with Support Vector Machines
Yamada et al. [7]	45	Game scores (Wii Fit Basic Step/Ski slalom)	Find correlations between game scores and a history of falls	Significant differences ( $p < .001$ ) and moderate correlations ( $r = 0.69$ ) identified

30

Table 1: Recent studies using WBB data to identify players at falling risk

31

32 In a previous article, we presented PDDanceCity, a city map exergame with the goal of providing dual-  
 33 tasked cognitive and physical rehabilitation [8]. The game is controlled with an array of six WBBs, which  
 34 we call Extended Balance Board (EBB) [9]. Thanks to its extended surface, EBB data can be used to  
 35 estimate the balance of the player both while standing and walking. We believe the data extracted from the  
 36 EBB could be used to estimate the balance and gait skills of the player in the background, without the need  
 37 to actively perform any specific test, or for any caregiver to be present.

38

39 In order to do so, this study aims to analyze the possibility of training a classifier to predict the falling risk of  
 40 a player based on EBB data collected in the background while playing PDDanceCity. This can be achieved  
 41 by attempting to predict the score of a standardized test that can be used to assess the falling risk. There  
 42 are several such tests to measure lower extremity strength, for example, the 30 second Sit-To-Stand test  
 43 (30CST) [10], which is part of the Fullerton Fitness Test Battery, and is fairly easy to administer. Evidence  
 44 suggests it is also a good screening test to estimate the falling risk and balance impairment in older adults  
 45 [11]. The 30CST classifies participants as subjects able or unable to maintain physical independence, a  
 46 binary prediction that could be achieved with a classifier algorithm.

47

48 Thus, the goal of this study is to determine the feasibility of classifying EBB-extracted data to perform a  
 49 binary prediction (player is able or unable to maintain physical independence) basing the result on a  
 50 prediction of the 30CST score. Data is collected while users are playing PDDanceCity to provide a very  
 51 simple background screening process determining whether the player may be at an increased falling risk.

52

53

## 54 Methods

55

56 PDDanceCity [8] is a labyrinth navigation exergame designed for dual-tasking rehabilitation. The goal of  
57 the game is to navigate a labyrinth, representing a city map to reach a goal, where only two-dimensional  
58 movements are possible (up, down, right and left). As an additional requirement, players are encouraged  
59 to reach the target with the least possible number of steps. In addition, they may be required to visit a given  
60 number of points of interest (for example a museum, monument or café) which may or may not be directly  
61 on the shortest path (Figure 1). The game offers a dual-tasking rehabilitation task, training visuospatial  
62 function, memory, balance, and physical coordination.

63

64 PDDanceCity is controlled with system consisting of an array of six WBBs, called EBB [9] (Figure 2). A  
65 controller receives all data from the WBBs and forwards it via a USB connection to a PC. Information sent  
66 through the USB interface contains the board identifier (ID), based on its MAC address, as well as the  
67 current value of each of its weighing sensors (four per WBB, for a total of 24). The refresh rate per board is  
68 20 Hz.

69

70 In order to use EBB data as a basis to control PDDanceCity, the center of mass  $\mathbf{com}(t)$  is calculated as  
71 follows. We define  $\mathbf{S}$  as the 6x4 matrix of sensor values (six WBB boards and four sensors per board), and  
72  $s_{i,j}(t)$  as the value of sensor  $(i, j)$  of  $\mathbf{S}$  in instant  $t$ . We define  $\mathbf{C}$  as the matrix of  $(x, y)$  coordinate vectors  $\mathbf{c}_{i,j}$   
73 assigned to each sensor (Figure 3), based on its position. We also define  $w(t)$  as the last total weight value  
74 calculated by all boards, that is, the weight of the player.  $\mathbf{com}(t)$  is calculated as the weight-normalized  
75 bidimensional projection of sensor values as:

76

$$77 \quad \mathbf{com}(t) = (com_x(t), com_y(t)) = \frac{1}{w(t)} \sum_{i=1}^6 \sum_{j=1}^4 (s_{i,j}(t) \mathbf{c}_{i,j})$$

78

79 This results in a set of two minus one to one values  $(com_x, com_y)$  which can be used to determine  
80 intentionality. To achieve this, we define a directional intention based on two conditions: the main directional

81 component must be equal to or greater than 0.5 in magnitude, and the other component must be equal to  
82 or lesser than 0.1 in magnitude. As an example, (0.1, 0.9) represents an upwards step, and  
83 (-0.8, 0.05) would represent a leftwards movement. Between each step, the player is always required to  
84 return to the center (both values lower than or equal to 0.1 in magnitude). Figure 4 represents two examples  
85 of this directional intention. We also define the instability factor  $if(t)$  as an approximation of the first order  
86 differential of  $com(t)$ . This parameter is a measure of how a player shifts their weight on the EBB. A very  
87 fast weight shifting, causing a high value of  $if(t)$ , would be an indicator of potential lack of balance (or loss  
88 thereof) among older adults who are not expected to move quickly. This is calculated as:

89

$$90 \quad if(t) = \sqrt{\frac{1}{2}(com_x(t) - com_x(t - 1))^2 + \frac{1}{2}(com_y(t) - com_y(t - 1))^2}$$

91

92 Where  $(t - 1)$  represents the value prior to the most recent one  $t$ . In this manner, when  $if(t)$  overcomes a  
93 certain threshold, a potential loss of balance may have occurred. For every level played, PDDanceCity  
94 stores a .xml file that includes the player's profile information, information about the level, steps taken and  
95 all values of  $com(t)$  and  $if(t)$ .

96

97 Finally, we extract a series of features based on  $com(t)$  and  $if(t)$ . These features are mostly related to  
98 average values, standard deviations and maxima and minima of  $com(t)$  under different circumstances, as  
99 well as the number of times that  $if(t)$  overcame different possible thresholds. In addition to these two  
100 elements, we also consider features related to the time intervals between steps, and the standard deviation  
101 of these intervals. A complete feature list is presented in Table 2. All features are calculated per playthrough,  
102 with no windowing. We used the Matlab software to calculate these features [12].

103

104 To evaluate our system, we recruited 16 participants (median age 73, 6 males) at a nursing home in  
105 Darmstadt, Germany. A computer was installed in a common room, connected to a television and the EBB  
106 (Figure 2). Participants were invited to play PDDanceCity as often as they desired for a period of two weeks.  
107 During the first session, nominal data (age and sex) was collected, and the 30CST was administered. All

108 sessions took place under observation of one of the authors, to ensure that no falls occurred. Otherwise,  
 109 the game sessions were unsupervised. We obtained approval of the ethics committee of the Technical  
 110 University of Darmstadt for this evaluation.

111  
 112

Table 2: System features and calculation

Features	Description	Calculation
$Com_{Avg_{Direction}}$	Average $com$ value for movements in each direction, where $n_{COM,Direction}$ represents the number of steps for each direction. Four two-dimensional features $(x, y)$ per playthrough	$\frac{\sum_{t=1}^{n_{COM,Direction}} com(t)}{n_{COM,Direction}}$ $Direction = Up \leftrightarrow com_y > 0.5,  com_x  < 0.1$ $Direction = Down \leftrightarrow com_y < -0.5,  com_x  < 0.1$ $Direction = Right \leftrightarrow com_x > 0.5,  com_y  < 0.1$ $Direction = Left \leftrightarrow com_x < -0.5,  com_y  < 0.1$
$Com_{Std_{Direction}}$	Standard deviation of $com$ , for each direction, as above. Eight features per playthrough	$\sqrt{\frac{\sum_{t=1}^{n_{COM,Direction}} (com_i(t) - Com_{Avg_{Direction,i}})^2}{n_{COM,i} - 1}}$ $i = x, y, Direction = Up, Down, Left, Right$
$Balance_{Up},$ $Balance_{Down}$	Average value of $com_y$ for all values where $com_y > 0$ (up) or $com_y < 0$ (down), where $n_{COM}$ is the number of $com$ samples. Two features per playthrough	$\frac{\sum_{t=1}^{n_{COM}} com_y(t)}{n_{COM}} : com_y > 0, \frac{\sum_{t=1}^{n_{COM}} com_y(t)}{n_{COM}} : com_y < 0$
$Balance_{Right},$ $Balance_{Left}$	Average value of $com_x$ for all values where $com_x > 0$ (right) or $com_x < 0$ (left). Two features per playthrough	$\frac{\sum_{t=1}^{n_{COM}} com_x(t)}{n_{COM}} : com_x > 0, \frac{\sum_{t=1}^{n_{COM}} com_x(t)}{n_{COM}} : com_x < 0$
$Avg_x, Avg_y$	Average value of $com_x$ and $com_y$ . Two features $(x, y)$ per playthrough	$\frac{\sum_{t=1}^{n_{COM}} com_x(t)}{n_{COM}}, \frac{\sum_{t=1}^{n_{COM}} com_y(t)}{n_{COM}}$
$Max_x, Max_y,$ $Min_x, Min_y$	Maximum and minimum value of $com_x$ and $com_y$ . Two features $(x, y)$ per playthrough	$Max (com_x(t), \forall t), Max (com_y(t), \forall t),$ $Min (com_x(t), \forall t), Max (com_y(t), \forall t)$
$Std_x, Std_y$	Standard deviation of $com_x$ and $com_y$ . Two features $(x, y)$ per playthrough	$\sqrt{\frac{\sum_{t=1}^{n_{COM}} (com_i(t) - Avg_i)^2}{n_{COM} - 1}}, i = x, y$
$If_{Avg}, If_{Max}$	Average $if(t)$ value and maximum for the whole playthrough. Two features per playthrough	$\frac{\sum_{t=1}^{n_{COM}} if(t)}{n_{COM}}, Max (if(t), \forall t)$
$If_{Threshold,i}$	Number of times $if(t) > i, i = [0.5, 1, 1.5, 2]$ . Normalized by the total number of samples. Four features per playthrough	$\frac{N (if(t) > i)}{n_{COM}}, i = 0.5, 1, 1.5, 2$
$If_{Sum_{Avg}},$ $If_{Sum_{Max}}$	Average value and maximum of the sum of the last 25 values of $if(t)$ for the whole playthrough. Two features per playthrough	$\frac{\sum_{t=1}^{n_{COM}} if_{Sum}(t)}{n_{COM}}, if_{Sum}(t) = \sum_{i=t-24}^t if(i),$ $Max (if_{Sum}(t), \forall t)$
$If_{Sum_{Overx}}$	Number of times $If_{Sum}(t) > i, i = [0.5, 1, 1.5, 2]$ . Normalized by total playthrough time. Four features per playthrough	$\frac{N (if_{Sum}(t) > i)}{n_{COM}}, i = 0.5, 1, 1.5, 2$
$Step_{Avg}$	Average time between steps, excluding the first step, defining $Step_{Time}(i)$ as the time in seconds in which step $i$ occurred, and $n_{Steps}$ as the total number of steps in the playthrough. One feature per playthrough	$\frac{\sum_{i=2}^{n_{Steps}} Step_{Time}(i) - Step_{Time}(i-1)}{n_{Steps}}$
$Step_{Std}$	Standard deviation of time between steps, excluding the first step. One feature per playthrough	$\sqrt{\frac{\sum_{i=2}^{n_{Steps}} (Step_{Time}(i) - Step_{Time}(i-1) - Step_{Avg})^2}{n_{Steps} - 1}}$

113

Table 2: System features and calculation

114

115 In total, these 16 participants played 87 levels of PDDanceCity during this period. For each level, a single  
 116 training instance was obtained. The data of 6 of these sessions had to be discarded due to data failure,  
 117 leaving 81 training instances for classification. To perform this classification, we attempted to classify if the  
 118 player's predicted 30CST score was above or below a cutoff score of 12 points, without using nominal data  
 119 (age and sex) as features. We refer to players classified above this cutoff as fit, and those under the cutoff  
 120 as not fit. This score was chosen to even out both groups, as eight participants had a 30CST score of 11  
 121 or lower. All classification tasks were performed using Weka [13].

122

123 **Results**

124

125 The best classification results are presented in Table 3 and the decision tree is depicted in Figure 6. A  
 126 comparison of different classification algorithms is presented in Figure 5. In all cases, we performed our  
 127 classification using ten-fold cross-validation. Results of a feature selection analysis (information gain  
 128 attribute evaluation) are included in Table 4. No features were excluded for classification.

129

130

Table 3: Best classification results

<b>Algorithm: Logistic Model Tree, accuracy 91.358%</b>	Correctly classified	Incorrectly classified	TP rate	FP rate	Precision	Recall	F	MCC	ROC area	PRC area
Not fit	29 (TP)	5 (FN)	0.853	0.043	0.935	0.853	0.892	0.823	0.940	0.946
Fit	45 (TN)	2 (FP)	0.957	0.147	0.900	0.957	0.928	0.823	0.940	0.930
Weighted average	74	7	0.914	0.103	0.915	0.914	0.913	0.823	0.940	0.936

131

Table 3: Best classification results using a Logistic Model Tree. TP = True Positive, FP = False Positive, F = F-measure, MCC = Matthews Correlation Coefficient, ROC = Receiver-Operating Characteristic Curve, PRC = Precision-Recall Curve

132

133 As a second potential scenario of analysis, we also aimed to predict the age- and sex-adjusted 30CST  
 134 cutoff scores. The resulting accuracy was very high (99%), but we suspect that to be due to overfitting to  
 135 age and sex because of our limited sample size, as the classifier did achieve 100% accuracy using  
 136 exclusively age and sex as features. If we remove these two features in this scenario, we achieve a  
 137 classification accuracy of 86% predicting the age- and sex-adjusted 30CST outcome. For this reason, we

138 believe that provided a large (and diverse) enough sample size of participants of a wide array of ages and  
 139 different degrees of fitness, it should be possible to predict the age- and sex-adjusted 30CST binary result  
 140 using the methods presented in this publication.

141

142

Table 4: Information gain attribute results

Feature	Information Gain ( $\Delta$ Entropy)
$Step_{Avg}$	0.486
$If_{SumOver0.5}$	0.321
$If_{SumOver1}$	0.241
$If_{Threshold,0.5}$	0.178
$Step_{Std}$	0.145
$If_{Avg}$	0.14
$If_{SumAvg}$	0.14
$If_{Max}$	0.138
$Balance_{Left}$	0.118

143

Table 4: Information gain attribute results based on delta entropy information gain

144

145 We complemented this classification with an analysis of the effect sizes of each feature between the fit and  
 146 not fit groups, measured on the basis of Hedges' g due to the low sample size and the disparity in standard  
 147 deviations. We also evaluated statistical significance using a t-test. These effect sizes are presented in  
 148 Table 5. Features related to the instability factor and the mean and standard deviation of the time between  
 149 steps seem to contain the most information related to the 30CST. In accordance with Cohen's rule of thumb  
 150 (0.2 is a small, 0.5 a medium and 0.8 a large effect size), the effect sizes of these features are large, with  
 151 the differences between the fit and not fit groups being in most cases very significant ( $p < 0.001$ ) or at least  
 152 significant ( $p < 0.05$ ).

153

154

155

156

157

158

159

160



161

Table 5: Effect size and statistical significance of features

Feature	Not fit-fit effect size (g)	Not fit-fit significance (p)	Feature	Not fit-fit effect size (g)	Not fit-fit significance (p)
<i>ComAvgUp,x</i>	0.3168	0.1998	<i>Avg<sub>y</sub></i>	-0.0056	0.9817
<i>ComAvgUp,y</i>	-0.4678	0.0595	<i>Max<sub>x</sub></i>	-0.3507	0.1528
<i>ComAvgDown,x</i>	0.2790	0.2692	<b><i>Max<sub>y</sub></i></b>	<b>-0.5279</b>	<b>0.0310</b>
<i>ComAvgDown,y</i>	-0.1171	0.6238	<i>Min<sub>x</sub></i>	0.3565	0.1472
<i>ComAvgRight,x</i>	-0.2328	0.3404	<i>Min<sub>y</sub></i>	-0.1036	0.6630
<i>ComAvgRight,y</i>	-0.4073	0.1028	<b><i>Std<sub>x</sub></i></b>	<b>-0.6665</b>	<b>0.0068</b>
<i>ComAvgLeft,x</i>	0.2661	0.2756	<i>Std<sub>y</sub></i>	-0.3789	0.1247
<i>ComAvgLeft,y</i>	-0.3234	0.1863	<b><i>If<sub>Avg</sub></i></b>	<b>-0.7478</b>	<b>0.0035</b>
<i>ComStdUp,x</i>	-0.0323	0.8966	<b><i>If<sub>Max</sub></i></b>	<b>-0.6337</b>	<b>0.0119</b>
<b><i>ComStdUp,y</i></b>	<b>-0.4913</b>	<b>0.0461</b>	<b><i>If<sub>Threshold,0.5</sub></i></b>	<b>-0.7452</b>	<b>0.0024</b>
<i>ComStdDown,x</i>	0.2234	0.3490	<i>If<sub>Threshold,1</sub></i>	-0.2411	0.2873
<i>ComStdDown,y</i>	0.4173	0.0860	<i>If<sub>Threshold,1.5</sub></i>	0	0
<i>ComStdRight,x</i>	0.0318	0.8977	<i>If<sub>Threshold,2</sub></i>	0	0
<i>ComStdRight,y</i>	0.0753	0.7621	<b><i>If<sub>SumAvg</sub></i></b>	<b>-0.7387</b>	<b>0.0038</b>
<i>ComStdLeft,x</i>	-0.3164	0.1959	<i>If<sub>SumMax</sub></i>	-0.2107	0.3938
<i>ComStdLeft,y</i>	-0.0237	0.9217	<b><i>If<sub>SumOver0.5</sub></i></b>	<b>-1.5261</b>	<b>&lt;0.0001</b>
<i>Balance<sub>Up</sub></i>	0.1598	0.5215	<b><i>If<sub>SumOver1</sub></i></b>	<b>-0.9196</b>	<b>0.0003</b>
<i>Balance<sub>Down</sub></i>	0.1623	0.4988	<i>If<sub>SumOver1.5</sub></i>	-0.2206	0.3477
<i>Balance<sub>Right</sub></i>	-0.3628	0.1429	<i>If<sub>SumOver2</sub></i>	-0.2062	0.3762
<b><i>Balance<sub>Left</sub></i></b>	<b>0.6306</b>	<b>0.0091</b>	<b><i>Step<sub>Avg</sub></i></b>	<b>1.2260</b>	<b>&lt;0.0001</b>
<i>Avg<sub>x</sub></i>	0.1925	0.4267	<b><i>Step<sub>Std</sub></i></b>	<b>0.8446</b>	<b>0.0020</b>

162

Table 5: Effect size and statistical significance of features. All values are presented as not fit vs. fit, meaning a negative effect size indicates the parameter has lower values in the not fit group. According to Cohen's rule,  $g > 0.8$  indicates a large effect size.  $p < 0.05$  denotes statistical significance

163

## 164 Discussion

165

166 Albeit the number of participants and training instances is somewhat limited, we obtained excellent  
 167 classification results. Generally, decision trees seem to provide the best performance in the proposed  
 168 classification task.

169

170 Our design also presents some technical limitations. At the moment, the WBBs send the data via Bluetooth,  
 171 which means they have to be manually connected for each play session. They also operate on batteries,  
 172 and when these are low the data received is not reliable anymore. Additionally, the EBB frame presents a  
 173 risk depending on how the EBB is placed in its surroundings: if it is not set against a wall behind it, a player  
 174 may fall when taking a step backwards. We aim to address these technical limitations in a future iteration

175 of the EBB by providing direct electrical supply to the WBBs, automating the Bluetooth synchronization  
176 process and building a complete enclosure around the EBB.

177

## 178 **Conclusions**

179

180 In spite of the aforementioned limitations, we believe our results suggest that the EBB, as an extension of  
181 the WBB, can be used to screen the elderly population for individuals with an increased risk of falling, as a  
182 basis to perform therapeutic and rehabilitation adjustments. Once the technical limitations of the EBB are  
183 addressed, and considering that participants played without supervision, a home (or, more generally,  
184 unsupervised) scenario seems feasible. In the future, we aim to extend our evaluation including features  
185 related to game performance, with the goal of performing a similar evaluation concerning cognition. This  
186 could be done, for example, on the basis of the Mini-Mental State Examination.

187

## 188 **Declarations**

189

### 190 **Ethics approval and consent to participate**

191 This work was submitted for consideration at the ethics committee of TU Darmstadt, case number EK09/20,  
192 approved on 17.03.2020. In accordance to the declaration of Helsinki, all participants willingly and freely  
193 took place in this evaluation, reading and signing an informed consent in which the goal of the study, their  
194 tasks, the information to be collected and the treatment of this information (pseudonymization) was  
195 presented in understandable language.

196

### 197 **Consent for publication**

198 Not applicable.

199

### 200 **Availability of data and materials**

201 Due to the nature of the data, the feature datasets of the presented classification are available from the  
202 corresponding author on reasonable request.

203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236

## Competing interests

The authors declare that they have no competing interests.

## Funding

This project did not receive specific funding. The authors are individually funded by their respective institutions.

## Author's Contributions

HB conducted the study at the nursing home with assistance from AGA, who designed the study and its goals. Data processing was performed by AGA with help from PNM and TT. AM and SG supervised the complete study, including improving the study design and protocol and assisting with the redaction of the manuscript. All authors have contributed to writing the manuscript and have reviewed it prior to submission.

## Acknowledgements

The authors would like to thank the administration and residents of the nursing home for their support and participation in the study.

## References

- [1] D. Aarsland, K. Brønnick, and T. Fladby, "Mild cognitive impairment in Parkinson's disease," *Current Neurology and Neuroscience Reports*, vol. 11, no. 4, pp. 371-378, 2011.
- [2] L. Z. Rubenstein, "Falls in older people: epidemiology, risk factors and strategies for prevention," *Age and Ageing*, vol. 35, no. Suppl 2, pp. 37-41, 2006.
- [3] S. Göbel, S. Hardy, V. Wendel, F. Mehm, and R. Steinmetz, "Serious games for health: personalized exergames," in *Proceedings of the 18th ACM international conference on Multimedia*, 2010, pp. 1663-1666.
- [4] R. Sun and J. J. Sosnoff, "Novel sensing technology in fall risk assessment in older adults: a systematic review," *BMC Geriatrics*. doi: 10.1186/s12877-018-0706-6 Available: <https://bmccgeriatr.biomedcentral.com/articles/10.1186/s12877-018-0706-6>
- [5] G. Mertes, G. Baldewijns, P.-J. Dingenen, T. Croonenborghs, and B. Vanrumste, "Automatic Fall Risk Estimation using the Nintendo Wii Balance Board," in *Healthinf*, 2015, pp. 75-81: Scitepress.
- [6] E. Bainbridge, S. Bevans, B. Keeley, and K. Oriel, "The effects of the Nintendo Wii Fit on community-dwelling older adults with perceived balance deficits: A pilot study," *Physical Occupational Therapy in Geriatrics*, vol. 29, no. 2, pp. 126-135, 2011.

237 [7] M. Yamada *et al.*, "The reliability and preliminary validity of game-based fall risk assessment in community-  
238 dwelling older adults," *Geriatric Nursing*, vol. 32, no. 3, pp. 188-194, 2011.

239 [8] A. Garcia-Agundez, A.-K. Folkerts, R. Konrad, P. Caserman, S. Göbel, and E. Kalbe, "PDDanceCity: An  
240 Exergame for Patients with Idiopathic Parkinson's Disease and Cognitive Impairment," presented at the  
241 Mensch und Computer 2017 - Tagungsband, Regensburg, Germany, Sep. 10-13, 2017.

242 [9] A. Garcia-Agundez, F. Baumgartl, F. Kendeffy, R. Konrad, H. Wunsch, and S. Göbel, "Development of a Wii  
243 Balance Board Array System for Exergames," in *Joint Int. Conf. Serious Games*, 2018, pp. 235-240: Springer.

244 [10] R. E. Rikli and C. J. Jones, *Senior Fitness Test Manual*. Human Kinetics, 2013.

245 [11] K. H. Cho, S. K. Bok, Y.-J. Kim, and S. L. Hwang, "Effect of lower limb strength on falls and balance of the  
246 elderly," *Annals of Rehabilitation Medicine*, vol. 36, no. 3, pp. 386-393, 2012.

247 [12] Mathworks, "MATLAB R2018a," ed, 2018.

248 [13] Machine Learning Group at the University of Waikato, "Weka 3," ed, 2020.

249

### **Titles to figures**

- Figure 1: PDDanceCity exergame
- Figure 2: PDDanceCity scenario setup
- Figure 3: EBB coordinate system
- Figure 4: EBB center of mass calculation
- Figure 5: Classification accuracies
- Figure 6: Decision tree

### **Legends to figures**

- Figure 1: PDDanceCity exergame
- Figure 2: PDDanceCity setup
- Figure 3: EBB coordinate system
- Figure 4: Calculation of the center of mass (top) and position of the feet (bottom) during a step forward ("a"), while standing on the center ("b") and during a step leftwards ("c")
- Figure 5: Classification accuracies
- Figure 6: Decision tree