



HAL
open science

A new deep learning-based method for the detection of gait events in children with gait disorders: Proof-of-concept and concurrent validity

Mathieu Lempereur, François Rousseau, Olivier Rémy-Néris, Christelle Pons,
Laetitia Houx, Gwenolé Quéllec, Sylvain Brochard

► To cite this version:

Mathieu Lempereur, François Rousseau, Olivier Rémy-Néris, Christelle Pons, Laetitia Houx, et al..
A new deep learning-based method for the detection of gait events in children with gait disorders: Proof-of-concept and concurrent validity. *Journal of Biomechanics*, 2020, 98, pp.109490 -. 10.1016/j.jbiomech.2019.109490 . hal-03488454

HAL Id: hal-03488454

<https://hal.science/hal-03488454>

Submitted on 21 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

A new deep learning-based method for the detection of gait events in children with gait disorders: proof-of-concept and concurrent validity

Mathieu Lempereur ^{*1,2,3}, François Rousseau^{1,4}, Olivier Rémy-Néris^{1,2,3}, Christelle Pons⁵,
Laetitia Houx^{1,2,3}, Gwenolé Quellec¹, and Sylvain Brochard^{1,2,3}

¹Laboratoire de Traitement de l'Information Médicale INSERM U1101, Brest, France

²Université de Bretagne Occidentale, Brest, France

³CHRU de Brest, Hôpital Morvan, service de médecine physique et de réadaptation, Brest, France

⁴IMT Atlantique, LaTIM U1101 INSERM, UBL, Brest, France

⁵Fondation ILDYS, Brest, France

Abstract

The stance and swing phases of the gait cycle are defined by foot strike (FS) and foot off (FO). Accurate determination of these events is thus an essential component of 3D motion recordings processing. Several methods have been developed for the automatic detection of these events (based on the heuristics of 3D marker position, velocity and acceleration), however the results may be inaccurate due to the high variability that is intrinsic to pathological gait. For this reason, gait events are still commonly determined manually, which is a tedious process. Here we propose a new application (*DeepEvent*) of a long short term memory recurrent neural network for the automatic detection of gait events. The 3D position and velocity of the markers on the heel, toe and lateral malleolus were used by the network to determine FS and FO. The method was developed from 10526 FS and 9375 FO from 226 children. *DeepEvent* predicted FS within 5.5 ms and FO within 10.7 ms of the gold standard (automatic determination using force platform data) and was more accurate than common heuristic marker trajectory-based methods proposed in the literature and another deep learning method.

A sensitivity analysis showed that *DeepEvent* mainly used the toe and heel markers (z-axis (longitudinal) position and velocity) at the beginning and end of gait cycle to predict FS, and the toe marker (x-axis (anterior/posterior) velocity and z-axis position and velocity) at around 60% of the gait cycle to predict FO.

Keywords

Gait, children, recurrent neural network, deep learning.

Word count

2025

*Corresponding author:

Address: CHRU de Brest, Hôpital Morvan, service de médecine physique et de réadaptation, 5 avenue Foch, 29609 Brest Cedex, France.
Telephone number: +33 (0)2-98-22-33-73 extension number 22426
Fax number: +33 (0)2-98-22-35-06
E-mail address: mathieu.lempereur@univ-brest.fr

1 Introduction

The current gold standard for the automatic detection of gait events in clinical practice involves the use of ground reaction forces measured by force platforms (Veilleux et al., 2016). However, pathological gait is highly variable and the use of gait aids, small steps or foot dragging during swing phase can produce false force thresholds, all of which may reduce the accuracy of the detection. To avoid these issues, gait events are therefore often determined manually, in particular for pediatric gait which is highly variable, however this process is time consuming and its accuracy and reliability strongly depend on operator expertise. To overcome these issues, computational methods of gait event detection, mostly based on kinematic data from opto-electronic systems, have been developed. A literature review (Bruening & Ridge, 2014) identified nine kinematic-based algorithms (De Asha et al., 2012; Desailly et al., 2009; Ghoussayni et al., 2004; Hreljac & Marshall, 2000; Hsue et al., 2007; Jasiewicz et al., 2006; O'Connor et al., 2007; Salazar-Torres, 2006; Zeni et al., 2008) that appeared promising for event detection in gait patterns such as equinus, slide/drag or steppage gait. The literature review and a recent study by Gonçalves et al. (2019) both revealed that the variables that result in the most accurate determination of events are sagittal resultant velocity (Ghoussayni et al., 2004), horizontal position (Desailly et al., 2009; Zeni et al., 2008), and vertical/horizontal acceleration (Hreljac & Marshall, 2000; Hsue et al., 2007) for FS, and horizontal position (Desailly et al., 2009; Zeni et al., 2008) and sagittal velocity (Ghoussayni et al., 2004) for FO. However, all those algorithms are limited by an accuracy of four frames (33.3 ms), and the fact that none can be used accurately across all types of gait pathologies. There is therefore a need for the development of learning-based approaches that do not require a priori information relative to the input data. Most studies that have used machine learning to study biomechanical parameters in patients with musculoskeletal and neuromuscular diseases have focused on the classification of pathological movements, prediction of the risk of development of certain diseases or the effect of an intervention, or the automatic recognition of activity for out-of-clinic monitoring (Halilaj et al., 2018). Recently, a method using deep neural networks to detect gait events in children with cerebral palsy determined FS and FO with respective errors of 18.3 ms and 12.5 ms (Kidziński et al., 2019). However, that method requires the use of a biomechanical model to compute the kinematic input data (3D joint angles using the conventional gait model described by Kadaba et al. (1990), 3D position and velocity of the markers, and 3D position, velocity and acceleration of the pelvis).

Recurrent neural networks, and more specifically long short term memory (LSTM) networks are among the most efficient in learning dynamics from time series (Hochreiter & Schmidhuber, 1997). Although LSTM networks are very popular (Schmidhuber (2015)), a specific type of analysis such as sensitivity analysis, Taylor decomposition or backward propagation, is necessary to understand the decision process used by such networks and thus to ensure appropriate interpretation of the output from the network, especially in the clinical context.

The primary aim of this study was to test the proof-of-concept of a new gait event detection method (*DeepEvent*) using an LSTM network. The second aim was to compare the accuracy of this new method with a set of different methods proposed in the literature and to determine if a resulting error of less than 10 ms (1 frame at 100 Hz) could be achieved. The third aim was to identify and interpret the parameters used in the decision process.

2 Method

2.1 Gait data

Data from 226 children who had undergone clinical 3D gait analysis were used. All parents provided written informed consent on a form that had been approved by the Brest CHRU. The children had the following pathologies: neurological disorders (65%, 76% of whom had cerebral palsy), neuromuscular disorders (11%), orthopedic disorders (9%) and other types of disorders (15%). All data were recorded using the same motion analysis system (Vicon MX, Oxford Metrics, UK) and four force platforms (Advanced Mechanical Technology, Inc., Watertown, MA, USA) in the same motion laboratory (Brest CHRU) between 2015 and 2019. The data collected by the 15 infrared cameras (sampling rate 100 Hz or 120 Hz) were synchronized with the ground reaction forces recorded by the force platforms (1000 Hz or 1200 Hz). The markers were placed according to the protocol by Kadaba et al. (1990). Marker trajectories and ground reactions forces were dual-pass filtered with a low-pass Butterworth filter at a cut-off frequency of 6 Hz. All participants walked independently, with no gait aids, at their natural speeds. An average of 5 trials was recorded for each child, yielding 1156 trials in total. All FS and FO events were manually identified by an engineer with 15 years' experience in gait analysis. The data included 10526 FS (5247 left and 5279 right) and 9375 FO (4654 left and 4721 right).

The data were randomly divided up such that: 60% were used for training (694 trials, 6303 FS (3139 left FS; 3164 right FS)), 5614 FO (2789 left FO; 2825 right FO)), 10% for validation (116 trials, 1039 FS (514 left FS; 525 right FS), 923 FO (457 left FO; 466 right FO)) and 30% for testing (346 trials, 3184 FS (1594 left FS; 1590 right FS), 2841 FO (1411 left FO; 1430 right FO)). These proportions are typically used in deep learning methodology (Crowther & Cox (2005)). Among the test data, 815 FS (446 left FS and 369 right FS) and 815 FO (369 left FO and 446 right FO) had previously been identified using the force platform data and were used as the gold standard method for comparison purposes. FS was determined as the first frame at which the vertical ground reaction force was greater than 10 N. FO was determined as the first frame at which the vertical ground reaction force was lower than 10 N.

2.2 Network architecture

3D position and velocity of the left and right heel, toe and lateral malleolus markers for each trial were used as input data for the *DeepEvent* recurrent network. A zero-padding technique was used to obtain the same number of frames for each trial (with a maximum duration of 1536 frames).

The neural network consisted of M bidirectional LSTM layers with N hidden units (Supplementary data) and was followed by a dropout layer. Finally, a one-time distributed dense layer was used to detect gait events. The 5 gait events (no event, left FS, right FS, left FO and right FO) were encoded using a one-hot encoding scheme. The parameters M and N were empirically chosen to maximize the performance of the network with the validation data set by an empirical approach. A sensitivity analysis was carried out to explain the neural network decisions by measuring the contribution of each input variable to the overall prediction (Simonyan et al., 2013).

DeepEvent was implemented using Keras (Chollet (2015)) in Python (Van Rossum & Drake (1995)), with cross-entropy as loss function and Adaptive Moment Estimation (Adam) for optimization on a Dell Workstation 7910 (CPU Xeon e5-2680-v4 2.4 GHz) with a Nvidia GeForce GTX 1080 Ti. *DeepEvent* is available on Github (github.com/LempereurMat/deepevent).

2.3 Validation: concurrent validity

The results of the *DeepEvent* method were compared to the position-, velocity- and acceleration-based algorithms proposed by Zeni et al. (2008), Ghoussayni et al. (2004) and Hsue et al. (2007) and the deep learning method proposed by Kidziński et al. (2019). The position-based algorithms reported by Zeni et al. (2008) and Desailly et al. (2009) detected FS with the same degree accuracy, however the method by Zeni et al. (2008) more accurately identified FO (Bruening & Ridge (2014)). The velocity-based method reported by Ghoussayni et al. (2004) identified gait events most accurately (Gonçalves et al. (2019)), and the use of horizontal acceleration resulted in a more accurate determination of FS and FO than the use of vertical acceleration (Hsue et al. (2007)).

The coordinate-based algorithm by Zeni et al. (2008) uses the anterior/posterior position of the heel and toe marker relative to the sacrum marker. The equation for the definition of FS is: $t_{FS} = \max(X_{Heel} - X_{Sacrum})$ and for FO is: $t_{FO} = \min(X_{Toe} - X_{Sacrum})$.

Ghoussayni et al. (2004) uses the sagittal plane resultant velocity and applies a gait speed-dependent threshold to detect FS: $FS_{Threshold} = 0.78 \times WalkSpeed$ and FO: $FO_{Threshold} = 0.66 \times WalkSpeed$.

Hsue et al. (2007) uses the peak horizontal deceleration of the heel marker to identify FS: $t_{FS} = \min(\frac{\partial^2 X_{Heel}}{\partial t^2})$ and the peak horizontal acceleration of toe marker to detect FO: $t_{FO} = \max(\frac{\partial^2 X_{Toe}}{\partial t^2})$.

Kidziński et al. (2019) uses the 3D joint angles using the conventional gait model described by Kadaba et al. (1990), 3D position and velocity of the markers, and 3D position, velocity and acceleration of the pelvis. Their network is built with 3 LSTM layers using a sequence of 128 frames. The results from *DeepEvent* were also compared to a modified version of the method by Kidziński et al. (2019), in which the number of frames was increased in order to obtain the same number of frames as in *DeepEvent* (1536).

Absolute differences in time were compared between the four algorithms in the literature, the modified Kidziński et al. (2019) algorithm, *DeepEvent*, and the force platform data.

3 Results

3.1 Proof-of-concept

Three bidirectional LSTM layers with 800 hidden units were chosen to maximize the performance of the network. With these parameters, the trained model had a validation cross-entropy of 0.0286 and an accuracy of 100%. The receiver operating characteristic analysis showed that the fully trained model achieved an area under the curve of 0.9955 for the left FS, 0.9971 for the right FS, 0.9958 for the left FO and 0.9955 for the right FO indicating that the model accurately determined FS and FO.

Computational times for training *DeepEvent*, the original and the modified Kidziński et al. (2019) methods were on average about 20, 8 and 42 minutes respectively.

3.2 Concurrent validity and parameters used in the decision process

DeepEvent predicted FS within 5.5 ms and within 10.7 ms of the gold standard, and was more accurate than the three methods from the literature, as well as the modified Kidziński et al. (2019) method (Table 1). These errors are acceptable since they are well below or just superior to the threshold of 10 ms.

136 The contribution of each type of input data to the decision process is presented in figures 1 and 2. The sensitivity
 137 analysis showed that *DeepEvent* used the input data between 0 and 20%, between 50 and 70% and between 90 and
 138 100% of the gait cycle. During these intervals, the contribution of each type of input data was between 0 and 0.04. Out-
 139 side of these intervals, the weights were closer to 0. These timings perfectly matched the definition of the beginning of
 140 swing phase at 60% of the gait cycle. The determination of the FS mainly involved the use of the z-axis (longitudinal)
 141 position and velocity of the toe and heel markers at the beginning and end of gait cycle (weight \approx 0.04). To predict
 142 FO, *DeepEvent* mainly used the x-axis (antero-posterior) velocity and z-axis position and velocity of the toe marker
 143 around 60% of the gait cycle (weight \approx 0.03).

	FS [CI]	FO [CI]
<i>DeepEvent</i>	5.5 [0.9;10.2]	10.7 (5.4;15.9)
Zeni et al. (2008)	39.7 [15.6;63.8]	13.7 [6.1;21.3]
Ghoussayni et al. (2004)	27.1 [11.4;43.0]	11.4 [5.9;16.9]
Hsue et al. (2007)	54.5 [25.7;83.3]	39.4 [26.9;51.9]
Kidziński et al. (2019)	8.6 [3.3;13.9]	22.7 [10.8;34.5]
Modified Kidziński et al. (2019)	6.3 [2;10.5]	11.2 [5.9;16.6]

Table 1: Absolute differences in milliseconds between *DeepEvent*, the methods in the literature, and the gold standard (automatic determination using force platform data). CI: confidence interval. Best results are highlighted in bold.

144 4 Discussion

145 The aim of this study was to test the proof-of-concept of a new gait event detection approach using an LSTM network.
 146 The results showed that *DeepEvent* determined FS and FO more accurately than three heuristic methods from the
 147 literature. Moreover, it was more accurate than the recently developed deep learning approach by Kidziński et al.
 148 (2019) that uses more input variables (kinematic outputs from conventional gait models). The high level of accuracy
 149 of *DeepEvent* means that it can be used in clinical practice for the automatic detection of the gait events with a
 150 very good inter-trial reliability. This study showed that the number of sequence frame (128 frames in Kidziński et al.
 151 (2019); 1536 in *DeepEvent*) is an important input parameter, since increasing the number of frames in the method by
 152 Kidziński et al. (2019) to the number used by *DeepEvent* improved the accuracy of the former method. The detection
 153 of FO was consistently less accurate than the detection of FS. This is not surprising, *DeepEvent* is a supervised
 154 LSTM-based approach. The accuracy of this type of algorithm is determined by the quality of the manual event
 155 annotation. In many cases, gait transitions, in particular FO, occur gradually and the identification of a single event
 156 frame is somewhat subjective, as reported by Kidziński et al. (2019). The location of the toe marker may also affect
 157 FO detection. The use of a hallux marker as in Stebbins et al. (2006) could increase FO detection accuracy but this
 158 marker placement is not commonly used in clinical gait analysis.

159 The sensitivity analysis demonstrated the importance of both the toe and heel markers for the accurate detection
 160 of FS and FO while the lateral malleolus marker did not appear to be essential. Equally, pelvic and other markers
 161 were unnecessary because the deep learning approach could determine gait direction due to the numerous cases in

162 the database of gait with direction along the anterior or posterior axis. In this study, all gait events were manually
163 identified by a single experienced engineer and the neural network was trained using this information. The results
164 now need to be externally validated by evaluating the robustness of this supervised technique using data from other
165 laboratories.

166 Acknowledgement

167 The authors would like to thank Stéphane Armand (Willy Taillard Laboratory of Kinesiology, Geneva University
168 Hospitals and Geneva University) for reviewing the manuscript and Johanna Robertson for language assistance.

169 Conflict of interest statement

170 None of the authors had any financial or personal conflict of interest with regard to this work.

171 References

- 172 Bruening, D. A. & Ridge, S. T. (2014). Automated event detection algorithms in pathological gait. *Gait & Posture*,
173 39(1), 472–477.
- 174 Chollet, F. (2015). Keras. <https://keras.io>.
- 175 Crowther, P. S. & Cox, R. J. (2005). A method for optimal division of data sets for use in neural networks. In Khosla,
176 R., Howlett, R. J., & Jain, L. C. (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems*, (pp.
177 1–7)., Berlin, Heidelberg. Springer Berlin Heidelberg.
- 178 De Asha, A. R., Robinson, M. A., & Barton, G. J. (2012). A marker based kinematic method of identifying initial
179 contact during gait suitable for use in real-time visual feedback applications. *Gait & Posture*, 36(3), 650–652.
- 180 Desailly, E., Daniel, Y., Sardain, P., & Lacouture, P. (2009). Foot contact event detection using kinematic data in
181 cerebral palsy children and normal adults gait. *Gait & Posture*, 29(1), 76–80.
- 182 Ghoussayni, S., Stevens, C., Durham, S., & Ewins, D. (2004). Assessment and validation of a simple automated
183 method for the detection of gait events and intervals. *Gait & Posture*, 20(3), 266–272.
- 184 Gonçalves, R. V., Fonseca, S. T., Araújo, P. A., Araújo, V. L., Barboza, T. M., Martins, G. A., & Mancini, M. C.
185 (2019). Identification of gait events in children with spastic cerebral palsy: comparison between the force plate and
186 algorithms. *Brazilian Journal of Physical Therapy*, *In press*.
- 187 Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J. L., Hastie, T. J., & Delp, S. L. (2018). Machine learning in human
188 movement biomechanics: Best practices, common pitfalls, and new opportunities. *Journal of Biomechanics*, 81,
189 1–11.
- 190 Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- 191 Hreljac, A. & Marshall, R. N. (2000). Algorithms to determine event timing during normal walking using kinematic
192 data. *Journal of Biomechanics*, 33(6), 783–786.
- 193 Hsue, B.-J., Miller, F., Su, F.-C., Henley, J., & Church, C. (2007). Gait timing event determination using kinematic
194 data for the toe walking children with cerebral palsy. *Journal of Biomechanics*, 40, S529.
- 195 Jasiewicz, J. M., Allum, J. H. J., Middleton, J. W., Barriskill, A., Condie, P., Purcell, B., & Li, R. C. T. (2006). Gait
196 event detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord injured
197 individuals. *Gait & Posture*, 24(4), 502–509.
- 198 Kadaba, M. P., Ramakrishnan, H. K., & Wootten, M. E. (1990). Measurement of lower extremity kinematics during
199 level walking. *Journal of Orthopaedic Research*, 8(3), 383–392.
- 200 Kidziński, L., Delp, S., & Schwartz, M. (2019). Automatic real-time gait event detection in children using deep neural
201 networks. *PLOS ONE*, 14(1), e0211466.
- 202 O’Connor, C. M., Thorpe, S. K., O’Malley, M. J., & Vaughan, C. L. (2007). Automatic detection of gait events using
203 kinematic data. *Gait & Posture*, 25(3), 469–474.

- 204 Salazar-Torres, J.-d.-J. (2006). Validity of an automated gait event detection algorithm in children with cerebral palsy
 205 and non-impaired children. *Gait & Posture*, *24*, S130–S131.
- 206 Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85 – 117.
- 207 Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification
 208 models and saliency maps. *CoRR*, *abs/1312.6034*.
- 209 Stebbins, J., Harrington, M., Thompson, N., Zavatsky, A., & Theologis, T. (2006). Repeatability of a model for
 210 measuring multi-segment foot kinematics in children. *Gait & Posture*, *23*(4), 401–410.
- 211 Van Rossum, G. & Drake, F. L. (1995). *Python tutorial*. Centrum voor Wiskunde en Informatica Amsterdam, The
 212 Netherlands.
- 213 Veilleux, L.-N., Raison, M., Rauch, F., Robert, M., & Ballaz, L. (2016). Agreement of spatio-temporal gait parameters
 214 between a vertical ground reaction force decomposition algorithm and a motion capture system. *Gait & Posture*,
 215 *43*, 257 – 264.
- 216 Zeni, J. A., Richards, J. G., & Higginson, J. S. (2008). Two simple methods for determining gait events during
 217 treadmill and overground walking using kinematic data. *Gait & Posture*, *27*(4), 710–714.

218 Supplementary data

219 The Long Short Term Memory (LSTM) is composed of four different gates that can change the information to the
 220 cell state in order to build a long-term memory (figure supplementary data). The first gate is called the "forget gate"
 221 (f_t) that decides for each element of the previous cell state vector C_{t-1} how much information to forget:

$$222 \quad f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

223 with σ the sigmoid function $\frac{1}{1+e^{-t}}$ and x_t the input data (3D position of the left and right lateral malleolus, 3D position
 224 of the left and right toe, 3D position of the left and right heel, 3D velocity of the left and right lateral malleolus, 3D
 225 velocity of the left and right toe and 3D velocity of the left and right heel). The second step concerns the "input
 226 gate" (i_t) and the "candidate gate" (g_t). The first one is used to update value in C_{t-1} while the second one is used
 227 to create a "candidate" that will be added to C_{t-1} . Then, the cell state is updated:

$$228 \quad i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$229 \quad g_t = \tanh(W_g[h_{t-1}, x_t] + b_c)$$

$$230 \quad C_t = f_t \odot C_{t-1} + g_t \odot i_t, \odot : \text{pointwise operator}$$

231 with \tanh a hyperbolic tangent function. Finally, the "output gate" (O_t) gives the output value (h_t):

$$232 \quad O_1 = \sigma(W_O[h_{t-1}, x_t] + b_O)$$

$$233 \quad O_2 = \tanh(C_t)$$

$$234 \quad h_t = O_1 \odot O_2$$

235 In the four gates, the values of the weight matrix W and the bias b were obtained during the training part of the
 236 recurrent neural network.