Adaptive Change-Point Detection for Studying Human Locomotion

Sylvain Jung, Laurent Oudre, Charles Truong, Eric Dorveaux, Louis Gorintin, Nicolas Vayatis, Damien Ricard

Adaptive Change-Point Detection for Studying Human Locomotion

Sylvain Jung[1,2,3,4], Laurent Oudre[1,2], Charles Truong[1,2],
Eric Dorveaux[4], Louis Gorintin[4], Nicolas Vayatis[1,2] and Damien Ricard[1,2,5,6]

***Abstract*— This paper presents an innovative method to analyze inertial signals recorded in a semi-controlled environment. It uses an adaptive and supervised change point detection procedure to decompose the signals into homogeneous segments, allowing a refined analysis of the successive phases within a gait protocol. Thanks to a training procedure, the algorithm can be applied to a wide range of protocols and handles different levels of granularity. The method is tested on a cohort of 15 healthy subjects performing a complex protocol composed of different activities and shows promising results for the automated and adaptive study of human gait and activity.**

***Clinical relevance*— A new approach to study human activity and locomotion in Free-Living Environments FLEs through an adaptive change-point detection which isolates homogeneous phases.**

I. Introduction and Background

Portable inertial sensors such as accelerometers, gyroscopes or Inertial Measurement Units (IMUs) are frequently used to analyze human activity. Among these activities, gait quantification is a major subject of interest for clinicians, as it can help to precociously detect the risk of falling or be applied in the context of longitudinal follow-up for degenerative diseases. Most of the published studies are led in clinical and controlled environments (laboratories, etc.), where efficient algorithms have been developed to extract from the raw data relevant features such as gait events, with precision as low as a few milliseconds [1].

However, these environments can induce a Hawthorne effect ("white coat" effect) on the patients and therefore bias the gait analysis process. A study has shown for instance that the variability of step duration in a free or semi-controlled environment is statistically different to the one measured in a controlled environment [2]. A wide range of new protocols, referred to as FLEs have therefore emerged to avoid this problem [3].

The study of locomotion in FLEs is more challenging, as the automatic computation of gait features at a micro level (stride times, step times...) is way more complex when no context information is available [4]. Recordings obtained in FLEs are generally analyzed in order to meet objectives set on a more general scale: activity classification [5], Energy Expenditure evaluation, falls' detection [6], etc. To the authors' knowledge, few or no studies based on the use of IMUs in FLEs have endeavored to propose a detailed macro analysis displayed in the shape of a visual simple caption.

In this article, a preliminary framework is presented for the study of IMUs signals recorded on healthy patients in FLEs. This framework is based on an adaptive change-point detection, that searches for significant changes in the time series at a given scale, i.e. instants where the subject changed his behaviour/activity. Instead of processing the signal in frames (as done in the vast majority of articles), the signal is processed at the scale of the output regimes. The used algorithm is adaptive in a sense that it uses a supervised procedure to learn the most appropriate scale for representing the time series [7]. This new approach makes it possible to identify specific changes that take place within the phases that are deemed to be stationary by other segmentation/classification procedures. Indeed, it highlights changes such as U-turns, turns or changes of pace within a single walking phase for instance, that are rarely identified in state-of-the-art activity classification systems. Such detailed segmentation enables the detection of relevant gait events (e.g. instabilities, hesitant steps, abrupt cadence changes...).

Adaptive gait breakpoint detection requires to specify two important parameters: the kind of changes we seek to detect and the scale of relevant breakpoints (which is linked to the sensitivity of the algorithm). In the proposed method, changes are detected in the spectrograms of the raw acceleration/angular velocity signals recorded on the lower back of healthy subjects. This data transformation has been successfully used in multiple publications as it enables to visualize activity changes but also speed changes or event slope changes [8], [9], [10]. As far as the scale is concerned, penalty learning approach described in [7] is adapted to the context of gait analysis. This method takes as input an annotated database of physiological signals where the relevant breakpoints have been annotated, and learns the adequate penalty parameter that allows to reproduce the strategy on new signals.

This study stands out from other approaches using activity classification processes to define at each instant of the signal which activity the recorded subject is performing. This way of analysis is widespread and well documented [3]. The aim here is to present a detailed, adaptive and structured visualization in order to better understand a complex gait

[1] Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, F-91190 Gif-sur-Yvette, France
[2] Université de Paris, CNRS, Centre Borelli, F-75005 Paris, France
[3] Université Sorbonne Paris Nord, L2TI, UR 3043, F-93430 Villetaneuse, France
[4] ENGIE Lab CRIGEN, F-93249 Stains, France
[5] Service de Neurologie, Service de Santé des Armées, Hôpital d'Instruction des Armées Percy, F-92190 Clamart, France
[6] Ecole du Val-de-Grâce, Ecole de Santé des Armées, F-75005 Paris, France

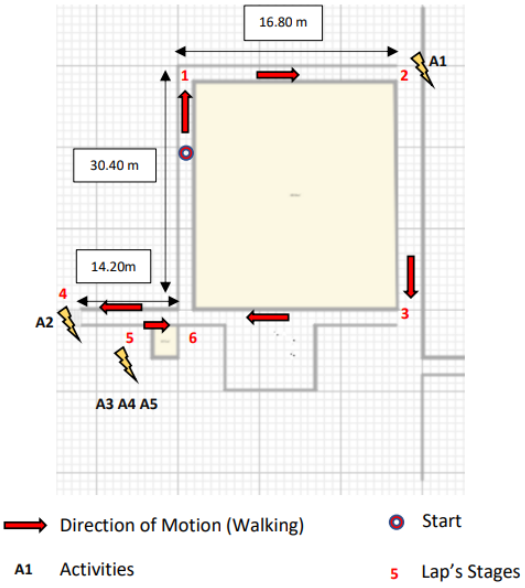


Fig. 1: Description of the semi-controlled protocol. Numbers displayed indicate the position of the subject during its path.

protocol.

II. Methods

A. Data

Fifteen healthy subjects (27.13±4.35 year-old, 7 men and 8 women) and one pathological patient (suffering from normal pressure hydrocephalus NPH) were measured on a semi-controlled protocol. Subjects were equipped with a Shimmer3 IMU on their lower back L5 (via the use of a belt strap) with sampling frequency F_s=100 Hz [3]. Subjects were asked to complete several laps of the Neurophysiology Department at Percy Hospital (a semi-controlled environment) for a total protocol duration of precisely 6 minutes.

The protocol, illustrated on Figure 1, contains several regimes that are either walking phases (denoted **W•**) or activities (denoted **A•**):

- (W1) Walk (1 → 2)
- (A1) Door Opening and 90 degrees Turn
- (W2) Walk (2 → 3)
- (W3) Walk (3 → 4)
- (A2) Going up 3 steps stairs U-turn and going down 3 steps stairs
- (A3) Leaning
- (A4) Standing Still
- (A5) Sitting Still
- (W4) Walk (5 → 6)
- (W5) Walk (6 → 1)

Transitions between two walking periods correspond to 90° Turn without door openings : they will be grouped in the same transition category since they are very similar.

B. Data Transformation

Two signals of interest are extracted from the raw data: the craniocaudal angular velocity (gCC) and the anteroposterior acceleration (aAP). These signals were chosen as they are directly influenced by breakpoints that can be observed during the execution of the protocol (beginning and end of gait, short activities, half turns and turns...). These signals are normalized (zero mean and unit variance) before being transformed in the time-frequency domain.

The magnitudes of the Short-Time Fourier Transform (STFT) coefficients of each of the signals aAP and gCC are computed. The STFT is performed with 300 samples per segment and an overlap of 290 samples. Only the 0 - 5 Hz frequency band, where phenomena of interest are contained, is kept. The obtained spectrograms are concatenated, thus providing $d = 28$ frequency bins per frame (14 per signal).

C. Annotations and Breakpoints

During each recording, an examiner (an expert able to identify breakpoints) has annotated all breakpoints described in Section II-A that correspond to transitions between walking phases/activities, activities/walking phases, activities/activities or walking phases/walking phases. These annotations consist in precise timestamps and will be used as ground truth breakpoints' labels in the following.

Intuitively, these breakpoints are of various nature, and in order to investigate their properties, the normalized mean-shift amplitude $\hat{\Delta}$ was computed for each breakpoints' type [9]. This quantity allows to investigate to what extend an annotated breakpoint τ separating two regimes *left/right* is *hard* to detect. Given a multivariate signal with d dimensions and an annotated breakpoint τ, the squared normalized mean-shift amplitude $\hat{\Delta}^2_\tau$ is defined as

$$\hat{\Delta}^2_\tau = \frac{1}{d}\sum_{i=1}^{d} \frac{\mu_{left,i} - \mu_{right,i}}{\frac{\sigma^2_{left,i}}{n_{left}} + \frac{\sigma^2_{right,i}}{n_{right}}}, \quad (1)$$

where $\mu_{left/right,i}$ and $\sigma^2_{left/right,i}$ are respectively the mean and variance of the signal on the left/right regime and dimension i, and $n_{left/right}$ the number of samples of left/right regimes. The higher $\hat{\Delta}_\tau$ is, the more different the two regimes are and the easier the breakpoints are to detect.

Table I displays the normalized mean-shift amplitude as well as additional information on all breakpoints considered in the article (type of regimes and regimes' length). It appears that transitions separating motion and static or "slow" phases tend to have higher $\hat{\Delta}$ values than those obtained for transitions dividing two movement phases. Transitions separating two static or "slow" phases present the lowest $\hat{\Delta}$ values (i.e. are the most difficult to detect) except for transition 5 which has one of the highest found value.

D. Adaptive Change-point Detection

Change-point detection is a central task in signal processing and consists in identifying the moments when a signal undergoes sudden changes [11]. Although unsupervised approaches for change-point detection are available [12] (likelihood ratio methods, subspace model methods...), parameters' settings for such a method is more complex than for supervised methods which enable automatic calibration.

Transition number	Type of breakpoints	Breakpoints characteristics			Accuracy results	
		$\hat{\Delta}$	Types of regimes	Regimes' length	Recall	Delta time (s)
1	W2→W3 W4→W5 W5→W1	6.68	Movement/Movement	Long/Long	0.73 ±0.09	1.87 ±0.14
2	W1→A1	7.20	Movement/Static-Slow	Long/Short	0.9 ±0.06	0.50 ±0.07
3	A1→W2	7.63	Static-Slow/Movement	Short/Long	0.96 ±0.05	0.55 ±0.16
4	W3→A2	11.04	Movement/Static-Slow	Long/Short	0.98 ±0.03	1.58 ±0.23
5	A2 → A3	10.13	Static-Slow/Static-Slow	Short/Short	0.90 ±0.10	0.88 ±0.36
6	A3 → A4	4.88	Static-Slow/Static-Slow	Short/Short	0.46 ±0.23	1.07 ±0.26
7	A4 → A5	4.79	Static-Slow/Static-Slow	Short/Short	0.83 ±0.09	1.11 ±0.42
8	A5 → W4	13.40	Static-Slow/Movement	Short/Long	0.94 ±0.05	1.16 ±0.05

TABLE I: Details of breakpoint's categories. The length of the regimes segmented by each kind of breakpoint are provided (short : $<$ 10s and long $>$ 10s) as well as the type of these regimes and the average number of times they occur in a lap. Accuracy results are also displayed. The recall rate corresponds to the ratio of the number of True Positive defined in section III by the number of annotated true breakpoints. Differences in seconds between a calculated breakpoint and the associated annotation correspond to Delta Times.

A supervised method is therefore used in this study. In this study's context, the changes to detect are transitions between activities of the protocol and it is assumed that those transitions correspond to mean shifts in the concatenated spectrograms (see Section II-B). Formally, let $y = \{y_1, y_2, \dots, y_n\}$ denote a $\mathbb{R}^d$-valued signal with n samples. For K change point indexes t_k ($1 < t_1 < t_2 < \cdots < t_K < n$), a common measure of approximation quality is the empirical quadratic risk:

$$R\left(y, \{t_k\}_k\right) := \sum_{k=0}^{K+1} \left(\sum_{t=t_k}^{t_{k+1}-1} \|y_t - \bar{y}_{t_k..t_{k+1}}\|^2 \right) \quad (2)$$

where $\bar{y}_{t_k..t_{k+1}}$ is the empirical mean of $y_{t_k}, \dots, y_{t_{k+1}-1}$ and $t_0 := 1$ and $t_{K+1} := n$ by convention. The risk (2) is simply the error when approximating y by a piecewise constant signal. The objective is to find the change points t_k that minimize this risk. When the number of breaks K is not known (which is the case in this article), the empirical quadratic risk is penalized with a linear penalty and the optimal breakpoints are

$$\hat{K}, \hat{t}_1, \dots, \hat{t}_K := \underset{K, t_1, \dots, t_K}{\arg\min} \left(\underbrace{R(y, \{t_k\}_{k=1}^K) + \beta K}_{R_\beta(y, \{t_k\})} \right) \quad (3)$$

The smoothing parameter β controls the trade-off between model complexity and model accuracy. (Note that $\hat{K}, \hat{t}_1, \dots, \hat{t}_K$ depend on β and y.) This value of β is critical, as it controls the sensitivity of the algorithm: large β will only detect strong breaks (change of activity for instance) while low β will also detect small breaks (change within the walking phases). For a fixed β, this discrete optimization problem is solved efficiently in linear time $\mathcal{O}(n)$ [13].

Instead of manually calibrating this parameter by trial and error, a supervised approach described in [7] is used in this article to learn the optimal parameter to detect the changes we are interested in. In a nutshell, this procedure takes as input a collection of N labeled signals $y^{(1)}, ..., y^{(N)}$, meaning that for each $y^{(i)}$, an expert manually provided the set of true change point indexes $\{t_k^{(i)}\}_k$. The optimal smoothing parameter, denoted $\hat{\beta}_{\text{opt}}$, is such that the risk of the true expert segmentation is closest to the one of the predicted segmentation, denoted $\{\hat{t}_k^{(i)}\}_k$:

$$\hat{\beta}_{\text{opt}} := \underset{\beta>0}{\arg\min} \frac{1}{N} \sum_{i=1}^{N} \left(R_\beta\left(y^{(i)}, \{t_k^{(i)}\}_k\right) - R_\beta\left(y^{(i)}, \{\hat{t}_k^{(i)}\}_k\right) \right). \quad (4)$$

Intuitively, the algorithm will search for the penalty β that allows to reproduce the annotations, by forcing the $\beta-$optimal solution $\{\hat{t}_k^{(i)}\}_k$ to be as close as possible to the ground truth partition $\{t_k^{(i)}\}_k$. The excess penalized risk is by definition a convex function w.r.t. β (affine function minus a concave function). We therefore used Brent's method as a convex optimization tool to minimize this component for each signal.

III. Results

In this section, two points are investigated: the ability of the algorithm to learn the relevant breakpoints in a gait signal, and the its ability to adapt to different annotation strategies.

A. Adaptive Change-point Detection

The change-point detection performances are measured with a 5-fold cross-validation: the optimal penalty value $\hat{\beta}_{\text{opt}}$ is learned using twelve gait signals and tested on the remaining three signals, and this operation is repeated five times. Each signal used for this cross-validation corresponds to an individual participant. This cross-validation allows to verify that the algorithm developed in this article can be used on new unseen data to apply the desired segmentation. A predicted change point is a True Positive (TP) if it is close to a true change point (within a margin of 3.5 seconds). This

margin corresponds to the maximum error for a transition observed in a human being from the recordings in this study. Recall is the ratio of the number of TPs to the number of true changes while precision is the ratio of the number of TPs to the number of predicted breakpoints. When a TP is detected, the difference in seconds between the calculated break and the associated annotation can be computed: this corresponds to the Delta Time metric.

On the complete data set, F1-score is equal to 0.81 ± 0.02 (precision: 0.82 ± 0.06, recall: 0.81 ± 0.04), meaning that this study's change-point detection procedure is accurate for most breakpoints, and also that there is no oversegmentation. Detailed performances are displayed on Table I. Breakpoints associated to transitions between walking phases and activities are especially well detected (transition types 2, 3, 4 and 8) with recalls above 0.9. Turns between two walking phases (transition type 1) are less precisely detected (recall 0.73). The poorest results are obtained for transitions 6 and 7 dividing the static activities (leaning, sitting, standing). This result is probably due to the fact that these transitions occur between two short regimes: transitions 6 and 7 are within two seconds from each other on average. Therefore, when a breakpoint is computed in their proximity, it is very challenging to define for which transition specifically it is computed. Their recalls results can therefore not be considered separately but jointly. In doing so, it appears that the recall rate for these transitions is the lowest than those obtained for all the other transitions (averaged to 0.64). These results emphasize the correlation between the observations made on $\widehat{\Delta}$ values (made in the II-C) and the difficulty of detecting these particular breakpoints.

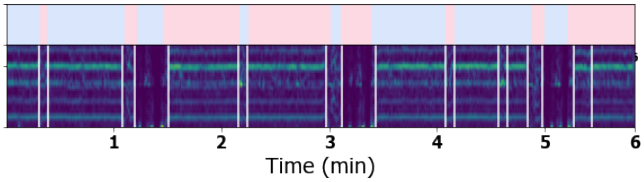
The largest Delta time error is obtained for Transition 4 (between the Walking phase W3 and the Stairs activity A2). This is due to the fact that the deceleration phase of the participants before the stairs appears well upstream of the stairs whereas the manual annotations often indicate them very close to the stairs. Nevertheless, it appears that these errors are quite limited.

B. Influence of the annotations

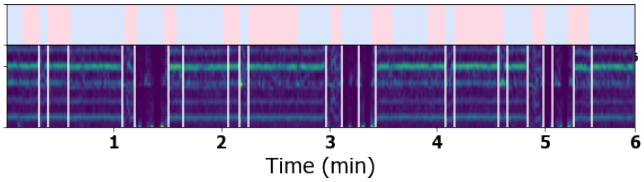
To understand the influence of the annotations on the segmentation results, three scenarios are imagined where the method is fed with different labels. These three scenarios are as follows:

- Scenario 1: only label transitions with a $\widehat{\Delta}$ above 7 (transitions 2, 3, 4, 5, and 8);
- Scenario 2: only label transitions with a $\widehat{\Delta}$ above 5 (transitions 1, 2, 3, 4, 5 and 8);
- Scenario 3: label all transitions.

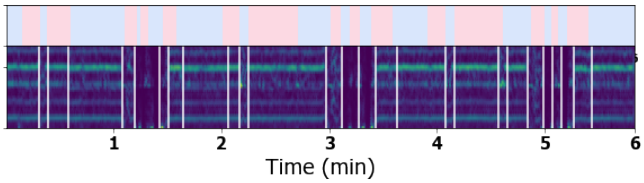
In Scenario 1, only the largest changes (with respect to $\widehat{\Delta}$) are provided by the expert. They correspond to transitions between activities. In Scenario 2, changes with medium amplitude (transitions between walking phases) are added to the previous set. Scenario 3, all transitions are labelled by the expert. This corresponds to the setting of the previous section. Results are illustrated on Figure 2.



(a) Scenario 1



(b) Scenario 2



(c) Scenario 3

Fig. 2: Evolution of the breakpoint's detection results depending on the annotations given as input. Displayed spectrograms are the concatenation of aAP and gCC spectrograms. White vertical lines correspond to the moments where the algorithm detects a rupture. Annotations given as input are displayed above each spectrogram (alternating colors).

They clearly show that the algorithm adapts well to the different scenarios and to the experimenters' annotations. Indeed, each time the annotations are reinforced with a new type of label, changes are induced on the detection of breakpoints and those changes fit to these input modifications. Moreover, it appears that this approach manages to adapt to the levels of granularity presented by the annotations structuring the scenarios. Thanks to this adaptability, the algorithm can reproduce the annotation strategy.

IV. Discussion

One of the interesting outcome of the proposed change-point detection method is that the obtained segmentation can be used to study individually each of the extracted homogeneous regimes. For instance, a set of features can be computed on each regime to characterize the whole timeline of the locomotion protocol. To illustrate this idea, two subjects were taken (one healthy subject and one patient with NPH), used this study's segmentation algorithm and computed on each regime the coefficients of variation of the Craniocaudal Angular Velocity (gCC) and the AnterioPosterior acceleration (aAP).

The obtained results are displayed on Figure 3 as color timelines (cold colors for small values and warm colors for large values). It appears that the segmented regimes in the patient with NPH are longer than in the healthy subject.

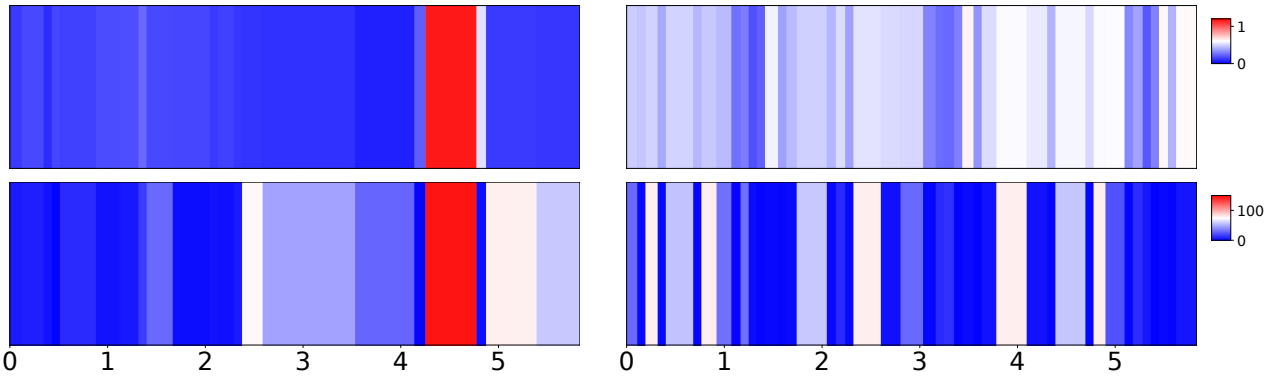


Fig. 3: Coefficients of variation for both the aAP (top) and gCC (bottom). A pathological subject (left) and a healthy subject (right) are considered. The coefficients are computed on each regime independently. On the x-axis, time is in minutes.

Moreover, these segments tend to be either very unstable or very stable in NPH subjects whereas the segments observed in healthy patients have CVs values close to the mean. One of the regimes of the patient with NPH is particularly unstable and corresponds to the moment when the patient has performed a slow U-turn. During this U turn, he showed a visible peak of instability (swaying to one side). This observation indicates that a clinician could quickly identify an abnormality in one of his patients' locomotion pattern and potentially associate it with an activity.

Thus, such a segmentation would provide an innovative feedback of a session in FLE because it is exhaustive, adaptable and precise. Indeed, this segmentation provides additional information that classification mechanisms do not. It is possible with this approach to go beyond the proportions of time spent on a particular activity. It is also possible to increase the volume of information provided to the practitioner by adding other prisms of evaluation than the one of stability (rhythm, sturdiness...). This method is also a quantitative feedback that can be adapted to specific annotations desired by the examiner (to target a sequence of activities for instance). Furthermore, it is a precise feedback that detects slight breaks reflecting changes that are not obvious but that reveal potential interpretations.

V. Conclusion

This method relying on an adapted learning of a penalty parameter through an optimization process enables a precise breakpoints' detection in acquisitions performed in semi-controlled environments. Thanks to this detection method, practitioners will be able to evaluate the stability of the segments found in their patients' recordings and to compare it with segments' stability of healthy patients. In this way, this method leads to a global visualization of a subject's activity at a macro level and to a beginning of an evaluation of the stationary phases detected at this macro level on long signals. This approach is a preamble to an innovative analysis of movements in completely FLEs. It will then be possible to optimize the β value for any kind of pathology and to automatically apply this method as a whole on new recordings from subjects suffering from the same pathology.

References

[1] T. Dot, F. Quijoux, L. Oudre, A. Vienne-Jumeau, A. Moreau, P.-P. Vidal, and D. Ricard, "Non-linear template-based approach for the study of locomotion," *Sensors*, vol. 20, no. 7, p. 1939, 2020.

[2] M. A. Brodie, M. J. Coppens, S. R. Lord, N. H. Lovell, Y. J. Gschwind, S. J. Redmond, M. B. Del Rosario, K. Wang, D. L. Sturnieks, M. Persiani *et al.*, "Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different," *Medical & biological engineering & computing*, vol. 54, no. 4, pp. 663–674, 2016.

[3] S. Jung, M. Michaud, L. Oudre, E. Dorveaux, L. Gorintin, N. Vayatis, and D. Ricard, "The use of inertial measurement units for the study of free living environment activity assessment: A literature review," *Sensors*, vol. 20, no. 19, p. 5625, 2020.

[4] F. A. Storm, K. Nair, A. J. Clarke, J. M. Van der Meulen, and C. Mazzà, "Free-living and laboratory gait characteristics in patients with multiple sclerosis," *PLoS One*, vol. 13, no. 5, p. e0196463, 2018.

[5] M. Nazarahari and H. Rouhani, "Detection of daily postures and walking modalities using a single chest-mounted tri-axial accelerometer," *Medical Engineering and Physics*, vol. 57, pp. 75–81, jul 2018.

[6] M. Nouredanesh and J. Tung, "IMU, sEMG, or their cross-correlation and temporal similarities: Which signal features detect lateral compensatory balance reactions more accurately?" *Computer Methods and Programs in Biomedicine*, vol. 182, p. 105003, dec 2019.

[7] C. Truong, L. Oudre, and N. Vayatis, "Penalty learning for changepoint detection," in *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 1569–1573.

[8] L. Oudre, J. Jakubowicz, P. Bianchi, and C. Simon, "Classification of periodic activities using the wasserstein distance," *IEEE transactions on biomedical engineering*, vol. 59, no. 6, pp. 1610–1619, 2012.

[9] C. Truong, "Détection de ruptures multiples–application aux signaux physiologiques." Ph.D. dissertation, Université Paris-Saclay, 2018.

[10] M. D. Nguyen, K.-R. Mun, D. Jung, J. Han, M. Park, J. Kim, and J. Kim, "Imu-based spectrogram approach with deep convolutional neural networks for gait classification," in *2020 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2020, pp. 1–6.

[11] L. Oudre, A. Lung-Yut-Fong, and P. Bianchi, "Segmentation of accelerometer signals recorded during continuous treadmill walking," in *2011 19th European Signal Processing Conference*. IEEE, 2011, pp. 1564–1568.

[12] S. Aminikhanghahi and D. J. Cook, "A survey of methods for time series change point detection," *Knowledge and information systems*, vol. 51, no. 2, pp. 339–367, 2017.

[13] R. Killick, P. Fearnhead, and I. A. Eckley, "Optimal detection of changepoints with a linear computational cost," *Journal of the American Statistical Association*, vol. 107, no. 500, pp. 1590–1598, 2012.