ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6. Available from http://www.i6doc.com/en/.

CDTW-based classification for Parkinson's Disease diagnosis

Nicolas Khoury, Ferhat Attal, Yacine Amirat, Abdelghani Chibani, and Samer Mohammed

LISSI Laboratory, Paris-Est Creteil University

Abstract. This paper presents a new classification approach for Parkinson's Disease (PD) diagnosis using Continuous Dynamic Time Warping (CDTW) technique and gait cycles data. These data are the vertical Ground Reaction Forces (vGRFs) recordings collected from eight force sensors placed in each shoe sole worn by each subject. The proposed approach exploits the principle of the repetition of gait cycle patterns to discriminate healthy subjects from PD subjects. The repetition of gait cycles is evaluated using the similarity of the time-series corresponding to stance phases estimated by applying the CDTW technique. The CDTW distances, extracted from gait cycles, are used as inputs of a binary classifier discriminating healthy subjects from PD subjects. Different classification methods are evaluated, including four supervised methods: K-Nearest Neighbours (K-NN), Decision Tree (DT), Random Forest (RF), and Support Vector Machines (SVM), and two unsupervised ones: Gaussian Mixture Model (GMM), and K-means. The proposed approach compares favorably with a classification based on standard features.

1 Introduction

Parkinson's disease (PD) affects the human central nervous system by destroying dopaminergic neurons which produce dopamine. Dopamine is a neurotransmitter that sends messages to the brain in order to control the human movement [1, 2, 3]. Thereby, most of PD patients present movement disorders causing walking disturbances [1, 2]. Several symptoms can be observed as: tremors, muscle stiffness, and changes in gait pattern [2]. The gait pattern is represented by the gait cycle which is mainly composed of two phases: stance phase and swing phase representing, respectively, 60 % and 40 % of the gait cycle. Usually, the walking of a healthy subject is characterized by a repetition of the gait cycle pattern, while that of PD subjects shows significant variations in the gait pattern from one gait cycle to another [4]. The repetition of gait cycles is evaluated using the similarity of the time-series corresponding to stance phases. Dynamic Time Warping (DTW) is a technique commonly used in the literature for time-series mining [5]. Continuous Dynamic Time Warping (CDTW), an improved version of DTW, is used in this paper to obtain a better resolution in the time-series matching, and therefore, a finer calculation of the similarity between gait cycles. The CDTW distances, extracted from gait cycles, are used as inputs of a binary classifier discriminating healthy subjects from PD subjects. This paper is organized as follows: section 2 is devoted to the datasets description and preprocessing phase. Section 3 presents the proposed approach. The performances of this approach are presented and discussed in section 4. Finally, a conclusion and some perspectives are given in the last section.

2 Datasets description and pre-processing

The three datasets used in this study are available on the PhysioNet website [6]. The first one is provided by Yogev et al. [7]; it includes the gait data of 29 PD subjects and 18 healthy ones. The second dataset, provided by Hausdorff et al. [8], contains the gait data of 29 PD and 25 healthy subjects. The third one, provided by Frenkel-Toledo et al. [9], includes the gait data of 35 PD subjects and 29 healthy ones. The gait data are the vertical Ground Reaction Forces (vGRFs) recordings collected from eight force sensors placed in each shoe sole worn by each subject. Each recording includes sixteen signals provided by the force sensors. In this study, the time-series corresponding to the sum of the 8 sensors outputs for each foot are used as reference to characterize the gait pattern of each subject. In the experiments, the subjects were asked to walk through a round trip walkway, leading therefore to the presence of outliers in the gait data. The recorded data in the turn-around phases are manually removed. Furthermore, to eliminate the start-up and end-up effects, the first and the last 20 seconds of the walking duration are also removed. Some fluctuations in the swing phase are observed, and lead to non-zero values in the vGRFs data; a 10-points median filter is used to remove these outliers.

3 Proposed diagnosis approach

The proposed approach follows three main steps: (1) Segmentation step: which consists of an automatic segmentation of each recorded gait time-series (left foot/right foot) taken separately, into a set of swing and stance phases. Only stance phases are considered in this study since vGRFs values in swing phases are equal to zero; (2) Feature extraction step: this step is achieved using the Continuous Dynamic Time Warping (CDTW) technique to evaluate for each foot the similarity between each stance phase and each of the other ones. Two CDTW distances vectors characterizing this similarity are then extracted from the segments obtained in the segmentation step; a distances vector is associated to each foot. By using the mean and the STD of each CDTW distances vector, four features are thus considered in the classification step; (3) Classification step: the four extracted features are used as inputs of the classifier. Different classification methods are evaluated, including four supervised methods: K-nearest neighbors (K-NN), Decision Tree (DT), Random Forest (RF), and Support Vector Machines (SVM), and two unsupervised ones: Gaussian Mixture Model (GMM), and K-means.

3.1 Continuous Dynamic Time Warping (CDTW)

This subsection describes the CDTW model used to calculate the similarity between gait cycles. Let $B_x = \{Q_x(t), t = 1, ..., T_x\}$ and $B_y = \{Q_y(t), t =$

ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6. Available from http://www.i6doc.com/en/.

$1, ..., T_y\}$ be 2-D curves, and $\phi = [\varphi_x(t), \varphi_y(t)]^T$ the warping (correspondence) map between B_x and B_y, with $Q_x(\varphi_x(t)) \in B_x$ corresponds to $Q_y(\varphi_y(t)) \in B_y$, for $t \in (1, ..., T)$. The distance between B_x and B_y is defined as follows:

$$\begin{aligned} & Dist(B_x, B_y) = \sum_{t=2}^{T} d((Q_{x(\varphi_x(t-1))}, Q_{y(\varphi_y(t-1))}), (Q_{x(\varphi_x(t))}, Q_{y(\varphi_y(t))})) \\ & = \sum_{t=2}^{T} \|Q_{x(\varphi_x(t))} Q_{y(\varphi_y(t))} - Q_{x(\varphi_x(t-1))} Q_{y(\varphi_y(t-1))}\|^2 \end{aligned} \tag{1}$$

The problem is to find the warping function $\phi = [\varphi_x(t), \varphi_y(t)]^T$ minimizing the distance in equation 1. The solution to this problem can be expressed as follows:

$$\phi = [\varphi_x, \varphi_y] = argmin_{\phi} Dist(B_x, B_y) \tag{2}$$

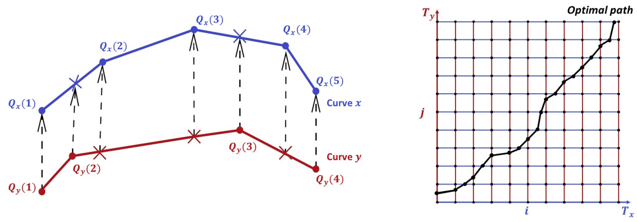
Let $Dist(t-1) = Dist(\phi(t-1))$ be the cumulated distance up to the $t-1$ matching, and $d((Q_{x(\varphi_x(t-1))}, Q_{y(\varphi_y(t-1))}), (Q_{x(\varphi_x(t))}, Q_{y(\varphi_y(t))}))$ the elementary distance, added by considering $Q_{x(\varphi_x(t))}$ corresponds to $Q_{y(\varphi_y(t))}$ knowing that $Q_{x(\varphi_x(t-1))}$ corresponds to $Q_{y(\varphi_y(t-1))}$. The cumulated distance $Dist(t)$ can then be expressed using the following recursion equation:

$$\begin{aligned} & Dist(t) = min_{\phi(t-1)}\{Dist(t-1) \\ & + d((Q_{x(\varphi_x(t-1))}, Q_{y(\varphi_y(t-1))}), (Q_{x(\varphi_x(t))}, Q_{y(\varphi_y(t))}))\} \end{aligned} \tag{3}$$

In the case of the DTW technique, a discrete solution is provided for the distortion function ϕ i.e. φ_x and φ_y take discrete values in $\{1, \ldots, T_x\}$ and $\{1, \ldots, T_y\}$, respectively. Therefore, a major limitation of this technique is the use of discrete points in the mapping of the two time-series. Furthermore, the DTW technique does not give the best 'optimal' warping path. To improve this technique, Munich et al. [10], proposed the Continuous DTW (CDTW) technique by applying the process of mapping in the continuous domain. In this case, the distortion function ϕ can take non-integer values as solution of the equation 3. Therefore, the CDTW technique allows the matching between sample point of one of the two time-series and another point between two samples of the other time-series. Unlike DTW, the warping path can pass between points on the grid vertices (figure 1.b). Thus, the recursion equation is similar to equation 3, under the following condition: if one of two points of the distortion function ϕ is an integer value, the second point of the distortion function ϕ can be a non-integer value. The intermediate matching point generation assumes a linear interpolation model between sample points, for more details [10]. Figure 1 illustrates the principle of the CDTW technique.

4 Results and discussion

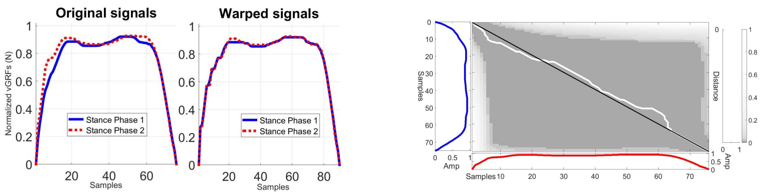
Figures 2.a and 2.c show the matching between the time-series of two right foot stance phases in the case of a healthy subject and a PD subject. One can



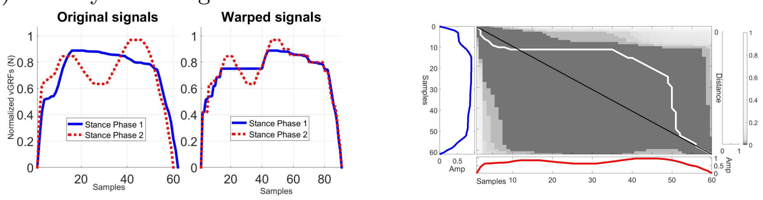
(a) Curve matching (b) Warping path

Fig. 1: Application of the Continous Dynamic Time Warping (CDTW)[10]

observe that the stance phases of the healthy subject are almost similar (figure 2.a), unlike those of the PD subject (2.c). The obtained 'optimal' warping paths are shown in white in figures 2.b and 2.d. The intensity of the gray color represents the value of the CDTW distance; the more the gray is darker the more the distance is higher, and vice-versa. The black lines represent the 'optimal' warping paths in the case of two identical stance phases. In the case of a healthy subject, the 'optimal' warping path is very close to that obtained with two identical stance phases (figures 2.b). Conversely, for a PD subject, the 'optimal' warping path is highly deformed and very far from that obtained in the case of two identical stance phases (figures 2.d).



(a) Healthy matching Stance Phases (b) Healthy Optimal path



(c) PD matching Stance Phases (d) PD Optimal path

Fig. 2: Results obtained with times-series of two right foot stance phases

In order to evaluate the performances of the proposed classification approach, a 10-fold cross validation is used.Before presenting the classification results, it is important to describe and set the parameters used in each classifier. The parameters of each standard classification method are chosen in such a way as to maximize their performances in terms of PD recognition. In the case of K-NN classifier, the number of neighbours is set by varying K from 2 to 10. The

ESANN 2018 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2018, i6doc.com publ., ISBN 978-287587047-6. Available from http://www.i6doc.com/en/.

optimal number of neighbours found for the Yogev's, Hausdorff's, and Frenkel-Toledo's datasets are, respectively, 3, 2, and 2. For the Decision Tree classifier, 'Cart' algorithm is used. For Random Forest classifier, the number of trees is varied from 10 to 500 and the optimal trees number found for the Yogev's, Hausdorff's, and Frenkel-Toledo's datasets are, respectively, 90, 200 and 300. For SVM classifier, a linear model is used in the case of Frenkel-Toledo's dataset, and a non-linear with polynomial kernel function of degree three in the case of the other two datasets. For the GMM classifier, a mixture of 2 diagonal Gaussians is used for the Frenkel-Toledo's dataset, and a mixture of 2 full Gaussians is used for the other two datasets. In the case of K-means classifier, the number of classes is equal to two (Healthy and PD subject classes). Table 1 shows the obtained performances in terms of accuracy rate and its standard deviation (STD). It can be noted that for the three datasets, SVM and K-NN classifiers, achieve the best performances. The accuracy rate obtained using these two classifiers range from 86 to 97 %. The accuracy rate range from 60 to 73 % using unsupervised classifiers.

		Supervised classifiers				Unsupervised classifiers	
Dataset	Performances	K-NN	CART	RF	SVM	K-means	GMM
Yogev et al.	Accuracy	92.88 %	82.96 %	89.35 %	93.57 %	67.79 %	65.58 %
	STD	2.03 %	3.99 %	2.62 %	2.61 %	1.10 %	7.23 %
Hausdorff et al.	Accuracy	97.52 %	88.82 %	90.02 %	95.03 %	65.29 %	73.97 %
	STD	1.04 %	3.09 %	1.98 %	1.85 %	2.13 %	4.90 %
Frenkel-Toledo et al.	Accuracy	86.02 %	80.02 %	82.59 %	87.32 %	60.47 %	67.19 %
	STD	3.09 %	6.44 %	4.92 %	2.99 %	1.48 %	8.65 %

Table 1: Accuracy rates and their STD

The performances of the proposed approach have been compared to those obtained using 5 features selected among 19 standard features[1] extracted from the swing, stance, and double stance phases. For this purpose, a wrapper approach, based on the random forest feature selection algorithm, is used [11]. Table 2 shows that the use of CDTW distance as distances features give the best performances. The accuracy rate improvement obtained with K-NN and SVM classifiers varies from 5 to 12 %.

5 Conclusion and future Work

This paper presents a new approach for the diagnosis of Parkinson's Disease (PD). This approach exploits the fact that usually the walking of healthy sub-

[1] Coefficients of Variation (CsV) in percentage (%) of the Swing Time of the left and right feet, CsV in duration (s) of the Swing Time of the Left and Right feet, CsV in duration of the Stride Time of the Left and Right feet, CsV of the Short Swing Time, CsV of the Long Swing Time, CsV of the Gait Asymmetry, Means in percentage (%) of the Swing Time of the Left and Right feet, Means in duration (s) of the Swing Time of the Left and Right feet, Means in duration (s) of the Stride Time of the Left and Right feet, Means in percentage of the Double Stance Time, Means of the Short Swing Time, Means of the Long Swing Time, and Means of the Gait Asymmetry.

Dataset	Features	Supervised classifiers				Unsupervised classifiers	
		K-NN	CART	RF	SVM	K-means	GMM
Yogev et al.	CDTW	$92.88 \pm 2.03\%$	$82.96 \pm 3.99\%$	$89.35 \pm 2.62\%$	$93.57 \pm 2.61\%$	$67.79 \pm 1.10\%$	$65.58 \pm 7.23\%$
	Standard	$85.39 \pm 3.49\%$	$82.10 \pm 3.25\%$	$85.65 \pm 3.23\%$	$85.02 \pm 4.26\%$	$63.72 \pm 4.20\%$	$64.77 \pm 12.64\%$
Hausdorff et al.	CDTW	$97.52 \pm 1.04\%$	$88.82 \pm 3.09\%$	$90.02 \pm 1.98\%$	$95.03 \pm 1.85\%$	$65.29 \pm 2.13\%$	$73.97 \pm 4.90\%$
	Standard	$91.39 \pm 2.83\%$	$84.82 \pm 3.66\%$	$88.77 \pm 2.30\%$	$87.70 \pm 3.26\%$	$55.12 \pm 3.59\%$	$65.12 \pm 11.08\%$
Frenkel-Toledo et al.	CDTW	$86.02 \pm 3.09\%$	$80.02 \pm 6.44\%$	$82.59 \pm 4.92\%$	$87.32 \pm 2.99\%$	$60.47 \pm 1.48\%$	$67.19 \pm 8.65\%$
	Standard	$81.49 \pm 4.68\%$	$79.01 \pm 3.72\%$	$81.98 \pm 4.19\%$	$80.23 \pm 4.80\%$	$57.19 \pm 3.98\%$	$65.31 \pm 12.12\%$

Table 2: Accuracy rates and their STD using the CDTW features (4 features) and the standard features (5 features)

ject is characterised by repetition of gait cycles, while that of PD ones show significant variations from one gait cycle to another. For this purpose, a measure of similarity between these cycles carried out using the Continuous Dynamic Time Warping (CDTW) technique is proposed. The obtained results showed an improvement in terms of accuracy rate for discriminating healthy subjects from PD subjects compared to the use of standard features. Ongoing work concerns the extension of the proposed approach to addressed the multidimensional similarity of the stance phase, i.e. the similarity between multiple stance phases.

References

[1] C.A.A. Davie, review of Parkinson's disease. *British medical bulletin*, 86(1):109-127, 2008.

[2] Parkinson's disease Foundation, (`http://parkinson.org/understanding-parkinsons`).

[3] C. Ahlrichs and M. Lawo, Parkinson's disease motor symptoms in machine learning: a review, *Health Inform. Int. J.*, 2(4), 2013.

[4] Y. Nancy Jane, H. Khanna Nehemiah and K. Arputharaj, A Q-backpropagated time delay neural network for diagnosing severity of gait disturbances in Parkinson's disease. J*ournal of biomedical informatics*, 60(C): 169-176, 2016.

[5] H. Sakoe and S. Chiba, Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, 26(1): 43-49, 1987.

[6] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark and H.E. Stanley, Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23): 215-220, 2000.

[7] G. Yogev, N. Giladi, C. Peretz, S. Springer, E. S. Simon, and J.M. Hausdorff, Dual tasking, gait rhythmicity, and Parkinson's disease: which aspects of gait are attention demanding?, *European journal of neuroscience*, Volume 22(5): 1248-1256, 2005.

[8] J.M. Hausdorff, J. Lowenthal, T. Herman, L. Gruendlinger, C. Peretz, N. Giladi, Rhythmic auditory stimulation modulates gait variability in Parkinson's disease, *European Journal of Neuroscience*, 26(8): 2369-2375, 2007.

[9] S. Frenkel-Toledo, N. Giladi, C. Peretz, T. Herman, L. Gruendlinger, J.M. Hausdorff, Treadmill walking as an external pacemaker to improve gait rhythm and stability in Parkinson's disease, *Movement Disorders*, 20(9): 1109-1114, 2005.

[10] M.E. Munich and P. Perona, Continuous dynamic time warping for translation-invariant curve alignment with applications to signature verification. *In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference*, 1: 108-115. IEEE, 1999.

[11] L. Breiman, J. Friedman, C.J. Stone, and R.A Olshen, Classification and Regression Trees, *CRC press*: Boca Raton, FL, USA, 1984.