Deep learning for Parkinson's disease: a case study on Freezing of Gait

Deep learning for Parkinson's disease: a case study on Freezing of Gait

Luigi Borzì, *Member, IEEE,* Marilena Varrecchia, *Member, IEEE,* Gabriella Olmo, *Senior Member, IEEE,*
Dept. of Control and Computing Engineering
Politecnico di Torino
Corso Duca degli Abruzzi 24, Torino, Italy
Email: {name.surname}@polito.it

***Abstract*—We propose a deep-learning method for feature extraction from gait data of Parkinson's disease patients. Our goal is to verify whether a fine classification of gait between similar groups can be achieved. To this end, we refer as a case study to the Freezing of Gait (FOG), and we measure gait data from two groups of patients, which exhibit (respectively, do not exhibit) this symptom. Wearable inertial sensors are employed, and data are collected during activities similar to those performed by patients during their daily living. Moreover, most patients are in *daily on* state, hence the two groups are difficult to classify, as their gait does not exhibit evident differences. Whereas classical Machine Learning methods are not sufficiently robust to perform such a fine classification, if they are fed with features extracted by means of a deep network, the results are satisfactory also when a large dataset is not available and data present a mild degree of heterogeneity.**

***Index Terms*—Deep learning, Parkinson's disease, freezing of gait, gait, smartphone, wearable device, inertial sensor**

I. Introduction

The clinical management of neurodegenerative disorders strongly benefits from data related to patient's activities in unsupervised environment. Recently, the Movement Disorder Society Task Force on Technology has published a roadmap for the development of patient-centered digital monitoring schemes, and their integration in the clinical practice [1]. Our activities on Parkinson's Disease (PD) are well in line with this roadmap. We want to use low-cost, commonly available instrumentation to seamlessly collect data from PD patients during their Activities of Daily Living (ADL). This can enable better follow up and clinical management of each single person, producing a sort of electronic diary tuned on each patient's requirements and conditions.

In this paper, we focus on Freezing of Gait (FOG), a very disturbing symptom that occurs in intermediate/advanced stages of the PD, related to fluctuating response to L-Dopa. FOG is very difficult to appreciate in the outpatient department due to its episodic nature. Moreover, it generally occurs when the patient is in *off* state, i.e., when the L-Dopa effects have weared off. We want to answer this question: is it possible to classify patients subject to FOG (FOG^+) using inertial data related to their gait during ADL, and in their *daily on* state, i.e. under the effect of prescribed medications? Is there some subtle difference between the pace of FOG^+ and FOG^- patients? Actually, gait impairments is known to be different in FOG^+ and FOG^- subjects [2], [3], and are related to the disease progression [4]. However, such differences may be difficult to catch, especially if the gait is measured on short time windows under unsupervised conditions, the patients are in *daily on* state, and the clinical disease progression (measured e.g. via the Hoehn Yahr scale) is similar between the two groups. On the other hand, the capability to perform a fine classification between FOG^+ and FOG^- can help revealing slight motor fluctuation during ADL, and also other activities and symptoms could be similarly monitored in order to achieve a global perspective of the patients status. Moreover, FOG impairs control of posture, increasing the risk of falls [5]; therefore, controlling this symptom can indirecty yield a prediction of fall risk.

First of all, we have faced the problem at hand by means of classical Machine Learning (ML) methods. We have identified a set of features already addressed in literature [6]–[8], and used them to feed several ML algorithms. Even if such features turned out to be significant, they revealed not sufficiently robust to perform a fine classification between FOG^+ and FOG^-. Hence, we propose an approach to extract features in a fully-automatic manner, in order to perform classification when a large dataset is not available and data present a mild degree of heterogeneity. The approach is robust: it ensures repeatability, as adding new patients does non entail a new feature computation, and good generalization potential.

II. FFT and CWT

Fast Fourier Transform (FFT) and Continuous Wavelet Transform (CWT), are largely employed tools to study aspects related to PD. FFT spectrum is mainly employed as a source for feature extraction [9], [10], while CWT is often exploited to perform gait analysis, as it enables easy detection of heel-strike and toe-off events with sound precision [11]. The method proposed in [11] relies on the CWT for step detection in both PD subjects and elderly people, using an inertial sensor placed on the lower back. In [4], walking parameters are extracted in free-living environment by exploiting wavelets.

The authors thank the staff of the University Hospital *Città della Salute e della Scienza*, Regional Reference Center for Parkinsons Disease and Movement Disorders, Turin (Italy) for the fruitful cooperation in performing the measures on PD patients.

FOG episodes were recently investigated with a convolutional neural network (CNN) in [12]. In order to detect FOG episodes in free-living conditions, authors collect accelometer and gyroscopic data from a smartphone stored in PD patient's trouser pocket. Then, signals are arranged in 2D images and input to the CNN. To the best of our knowledge there is no approach that exploits images representing FFT spectrum or CWT spectrogram to perform a gait assessment of PD patients during ADL.

III. Materials and Methods

The data employed for our experiment have been collected at the University Hospital *Città della Salute e della Scienza*, Turin (Italy), which hosts the Regional Reference Center for Parkinsons Disease and Movement Disorders. A total of 26 PD patients were recruited, and divided into FOG$^+$ and FOG$^-$ groups depending on their clinical records. Inclusion criteria were a clinical diagnosis of PD with motor symptoms, no major comorbidities, no significant vision or cognitive impairments, and no mobility aid. All subjects were in *daily on* state. The clinical characteristics of both groups are reported in Tab. I. The study has been conducted in accordance with the Declaration of Helsinki and approved by the local Ethics Committee.

Group	Subjects (male)	Mean age years $\pm$ SD	Years of disease	H&Y
FOG$^+$	14(7)	74.8 ± 6.4	8.6 $\pm$7.0	2.8 ± 0.6
FOG$^-$	12(8)	65.7 ± 10.7	7.1 ± 4.9	2.1 ± 0.2

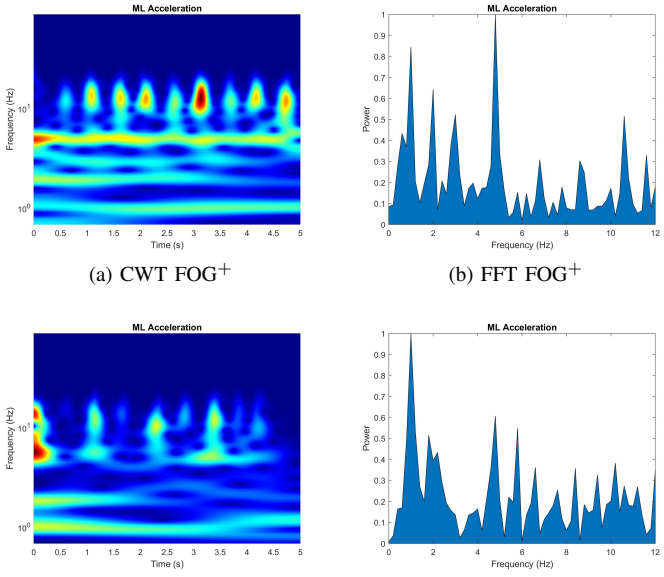
TABLE I: PD characteristics.

Data were acquired during the scheduled clinic assessment in the outpatient department in semi-supervised conditions, so as to mimic ADLs in a typical home environment. During each session, accelerometer (3-axis) and gyroscopic (3-axis) signals were recorded by means of a smartphone (Samsung S5 mini) locked in a Velcro belt tied around participant's waist. This location is close to body mass center, and guarantees the maximum patient's comfort [1]. Inertial sensor data were collected by means of the Android free App *SensorLog* [13] and locally stored. Then, once exported in *CSV* format, data were processed offline using MATLAB R2018b.

A. Data preprocessing

Inertial data were detrended and filtered with a 2nd order zero-lag Butterworth low-pass filter with cutoff frequency 20 Hz [14]. In this study we only focus on gait, thus data windows related to walking activity were extracted from the inertial data. To overcome the dataset heterogeneity (in terms of walked distance, duration of windows, number of steps), we have extracted only central 5s from each detected walking bout. The total number of windows amounts to 220 per axis (82 FOG$^+$ and 138 FOG$^-$).

Then, we have turned 1-D signals into 2-D images in order to feed a CNN. To this end, two types of image datasets were built up using CWT (Morlet analytic wavelet), and FFT spectrograms (Fig. 1). As CWT and FFT hold somewhat related information, in order to understand if and how results are affected by the kind of input images, the CNN performance has been evaluated on both datasets.



(a) CWT FOG$^+$ (b) FFT FOG$^+$

(c) CWT FOG$^-$ (d) FFT FOG$^-$

Fig. 1: CWT and FFT comparison. (a) and (b) ((c) and (d) respectively) refer to the same window of a FOG$^+$ (FOG$^-$) patient. Displayed component: M-L acceleration.

B. Deep learning for feature extraction

Gait features have been selected and extracted exploiting pre-trained CNNs, as the available dataset is not large enough to build up a net from scratch. Two types of CNNs, namely *AlexNet* [15] and *VGG-16* [16], exhibit a satisfactory trade-off between performance and computational time. In order to determine the best architecture for our classification purposes, we have tested the CNNs on all components of both accelerometer and gyroscope data. Furthermore, we have tuned the layer to extract features from; finally, for each component, we have selected the combination (i.e. CNN type and layer) returning the best accuracy. The extracted features were fed to a SVM with linear kernel. Since six outputs are provided by the CNN, i.e. one for each acceleration and gyroscope component, we computed the final class as the weighted mean of classificator outcomes, with each weight fitting the corresponding SVM accuracy. In order to limit computational complexity, we have included the minimum set of components ensuring the best final performance. To this purpose, a misclassification rate minimization was performed. The most meaningful dimensions for data classification turned out to be the mediolateral (M-L) acceleration and the angular velocity signal around the vertical (V) direction, irrespective of the type of image employed (CWT or FFT). The final configurations are described in Tab. II for both CWT and FFT scenarios.

Sensor	Net	Layer	Accuracy	# of features computed
		CWT		
M-L Acc.	$VGG-16$	ReLU 5.3	89.0%	100352
V Gyr.	$AlexNet$	FC7	91.1%	4096
		FFT		
M-L Acc.	$VGG-16$	ReLU 5.3	89.5%	100352
V Gyr.	$VGG-16$	FC6	86.8%	4096

TABLE II: Selected CNNs architecture. FC: *Fully Connected layer*. ReLU: *Rectified Linear Unit*.

The fusion of layouts reported in Tab. II achieves a global accuracy (in terms of 5-fold validation) of 93.6% and 90.0% for CWT and FFT respectively. In order to assess robustness, we have also carried out a Leave-one-patient-out (LOPO) validation. As expected, the performance slightly impairs, while maintaining reasonable values of 82.3% (with CWT) and 84.5% (FFT). Results are reported in Fig. 2 (labeled M1).

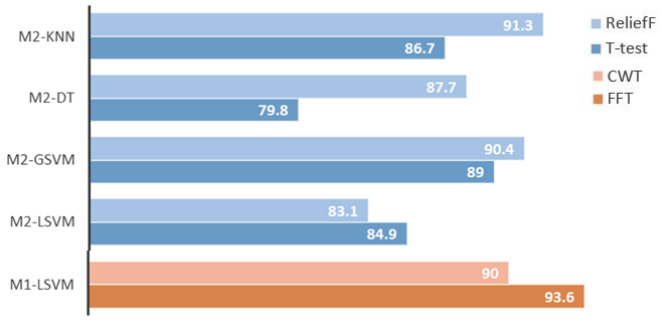
C. Manual feature selection

A wide set of features has been selected by means of a thorough literature analysis. The inclusion criterion is the recognized capability to describe gait and postural control in healthy, frail, faller, and/or PD freezer subjects. A total number of 68 features has been identified, regarding all components of both acceleration and angular velocity, and related to movement intensity, gait smoothness, symmetry, regularity, step and stride frequency. Once being extracted from each walking window, the feature mean values evaluated on FOG^+ and FOG^- patients have been compared. After having assessed the data normality using the Lilliefore test, the F-test on variances followed by T-test on mean values were performed for those features exhibiting a normal distribution, whereas Kolmogorov-Smirnov test followed by Wilcoxon test on median values were addressed in case of non-normal distribution.

Two different approaches have been validated for feature selection. The first one consisted in selecting features with different mean (median) values in FOG^+ and FOG^- sets (M2-A), while the second one is based on the ReliefF algorithm for assessing predictor relevance (M2-B). The two obtained training sets (39 and 25 features respectively, listed in Tab. III) have been input to several ML algorithms, and a 5-fold cross validation was performed in order to avoid overfitting. Moreover, LOPO validation was executed in order to test the classifier robustness. Results of both validation processes for different ML algorithms are reported in Fig. 2. For the sake of brevity, in Fig. 2b only the best classifier performance is displayed (SVM linear with Bayesian hyperparameters optimization). The dramatic performance impairment of M2 in LOPO validation can be explained by the different data management of the two validation algorithms. Being the dataset composed of more than one window from each patient, and since windows for training and test are randomly selected by the cross-validation algorithm, it is likely that some data windows of the patient under test are also used in the training phase; this reduces the reliability of the results. LOPO avoids this shortcoming, and provides a sensible validation of the classifier robustness and generalization capability.

Source	Extracted features	FS M2-A (p)	FS M2-B
[9], [17], [18]	RMS	1:3,6 (<0.0002)	3,6
[19]	Range	1:3,6 (<0.0001)	3,6
[18]	Jerk	1:3 (<0.0001)	2,3,6
[18]	Normalyzed jerk	1:3 (<0.0002)	2,3
[20]	Spectral Entropy	1:3,6 (<0.0002)	3
[6], [20]	Spectrum Peaks	1:4,6 (<0.0148)	3,6
[20]	Normalyzed Spectrum Peaks	1,3,6 (<0.004)	3,4,6
[21]	Harmonic Index	2 (<0.0001)	ns
[22]	Low Power Frequency	2:6 (<0.015)	ns
[23], [7]	Step variability	1 (<0.0001)	1
[6], [22], [23]	Stride variability	1 (<0.0001)	ns
[22], [23]	Symmetry	1 (0.0047)	1
[23]	Step time	ns	1
[22], [23]	Stride time	ns	1
[6], [20], [7]	Dominant frequency	ns	1,3,5,6

TABLE III: Selected features, M2-A and M2-B methods. (1:3) and (4:6) represent V, M-Ll and antero-posterior acceleration and angular velocity. Features marked with *ns* were not selected.



(a) 5-fold validation.

(b) LOPO validation.

Fig. 2: Validation results of several classifiers. Fig. 2b refers to Linear SVM classificator.

IV. Discussion

The classification results obtained with CNN feature extraction are promising, and outperform all classifers fed with manually extracted features in the LOPO validation. In fact,

different of classic ML approaches, features extracted and selected by means of CNNs are able to capture subtle differences in similar datasets. We believe that this result is not due to improper feature selection in M2, since these features are indeed significant in distinguishing the two samples (see Tab. III) and have yielded high accuracy in the cross validation phase (Fig. 2). Moreover, the classification performance is not solely related to the dimension of the feature set, and may even significantly impair with the addition of features, particularly if a small training set is used [24]. An automated feature extraction and selection substantially reduces the user-dependency in the feature selection task, and errors in their implementation. Moreover, the process is domain-independent and could be used for a variety of classification problems. The input data preprocessing is much less expensive for M1 compared to M2. On the other hand, computational times and memory requirements for the training phase of M1 are significantly greater than those of M2. Nevertheless, once the net is trained, M1 processing times are no longer very expensive (the feature extraction from a 3.5 min-long walking signal last about 1.5s).

V. Conclusions and future work

In this paper, we employ deep-learning for feature extraction from gait data of Parkinson's disease patients. Our objective is to employ wearable inertial sensors and to collect gait data from PD patients during their activities of daily living. We have shown that our method is effective in classifying patients that exhibit Freezing of Gait from those which do not, also when a large dataset is not available and data present a mild degree of heterogeneity. Such an approach could be employed to assess mild motor fluctuations during ADL, not only related to gait, but also to other activities such as standing, turning, postural instability for fall risk assessment. Future developments will be in the direction of exploiting deep learning to achieve a patient-centered follow-up of people affected by PD (the *electronic diary*), as well by other neuro-degenerative disorders.

References

[1] Alberto J Espay and et.al, "A Roadmap for Implementation of Patient-Centered Digital Outcome Measures in Parkinson's disease Obtained Using Mobile Health Technologies," *Movement Disorders Clinical Practice (in press)*, pp. 1–7, 2019.

[2] "Increased foot strike variability in parkinson's disease patients with freezing of gait," *Parkinsonism & Related Disorders*, vol. 53, pp. 58 – 63, 2018.

[3] Nellie Georgiou-Karistianis, Rachel Chee, Anna Murphy, Mary Danoudis, and Robert Iansek, "Gait freezing in Parkinson's disease and the stride length sequence effect interaction," *Brain*, vol. 132, no. 8, pp. 2151–2160, 05 2009.

[4] Silvia Del Din, Alan Godfrey, Brook Galna, Sue Lord, and Lynn Rochester, "Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length.," *Journal of neuroengineering and rehabilitation*, vol. 13, no. 1, pp. 46, 2016.

[5] Aner Weiss, Talia Herman, Nir Giladi, and Jeffrey M. Hausdorff, "Objective Assessment of Fall Risk in Parkinson's Disease Using a Body-Fixed Sensor Worn for 3 Days," *PLoS ONE*, vol. 9, no. 5, pp. e96675, may 2014.

[6] Aner Weiss, Talia Herman, Nir Giladi, and Jeffrey M. Hausdorff, "Objective assessment of fall risk in Parkinson's disease using a body-fixed sensor worn for 3 days," *PLoS ONE*, vol. 9, no. 5, 2014.

[7] Aner Weiss, Talia Herman, Nir Giladi, and Jeffrey M. Hausdorff, "New evidence for gait abnormalities among Parkinson's disease patients who suffer from freezing of gait: insights using a body-fixed sensor worn for 3 days," *Journal of Neural Transmission*, vol. 122, no. 3, pp. 403–410, mar 2015.

[8] J L Bellanca, K A Lowry, J M Vanswearingen, J S Brach, and M S Redfern, "Harmonic ratios : A quantification of step to step symmetry," *Journal of Biomechanics*, vol. 46, no. 4, pp. 828–831, 2013.

[9] B Huijben, "The effect of walking speed on quality of gait in older adults Gait & Posture The e ff ect of walking speed on quality of gait in older adults," *Gait & Posture*, vol. 65, no. January 2019, pp. 112–116, 2018.

[10] Aliénor Vienne, Rémi P. Barrois, Stéphane Buffat, Damien Ricard, and Pierre Paul Vidal, "Inertial sensors to assess gait quality in patients with neurological disorders: A systematic review of technical and analytical challenges," *Frontiers in Psychology*, vol. 8, no. MAY, pp. 1–12, 2017.

[11] Minh H. Pham and et.al, "Validation of a Step Detection Algorithm during Straight Walking and Turning in Patients with Parkinson's Disease and Older Adults Using an Inertial Measurement Unit at the Lower Back," *Frontiers in Neurology*, vol. 8, pp. 457, 2017.

[12] KimHan Byul, LeeHong Ji, LeeWoong Woo, KimSang Kyong, JeonHyo Seon, ParkHye Young, ShinChae Won, YiWon Jin, JeonBeomseok, and ParkKwang S., "Validation of Freezing-of-Gait Monitoring Using Smartphone," nov 2018.

[13] AAVV, "Sensorlog," 2014.

[14] Jin Wang, Ronghua Chen, Xiangping Sun, Mary F.H. She, and Yuchuan Wu, "Recognizing human daily activities from accelerometer signal," *Procedia Engineering*, vol. 15, pp. 1780–1786, jan 2011.

[15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, USA, 2012, pp. 1097–1105, Curran Associates Inc.

[16] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv 1409.1556*, 2014.

[17] Mark D. Latt, Hylton B. Menz, Victor S. Fung, and Stephen R. Lord, "Acceleration patterns of the head and pelvis during gait in older people with Parkinson's disease: A comparison of fallers and nonfallers," *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, vol. 64, no. 6, pp. 700–706, 2009.

[18] Mandy Miller Koop, Sarah J. Ozinga, Anson B. Rosenfeldt, and Jay L. Alberts, "Quantifying turning behavior and gait in parkinsons disease using mobile technology," *IBRO Reports*, vol. 5, pp. 10 – 16, 2018.

[19] Alejandro Galán-Mercant and Antonio I. Cuesta-Vargas, "Differences in trunk accelerometry between frail and non-frail elderly persons in functional tasks," *BMC Research Notes*, vol. 7, no. 1, pp. 1–9, 2014.

[20] Moacir Ponti, Patricia Bet, Caroline L. Oliveira, and Paula C. Castro, "Better than counting seconds: Identifying fallers among healthy elderly using fusion of accelerometer features and dual-task Timed Up and Go," *PLoS ONE*, vol. 12, no. 4, pp. 1–21, 2017.

[21] C. J.C. Lamoth, P. J. Beek, and O. G. Meijer, "Pelvis-thorax coordination in the transverse plane during gait," *Gait and Posture*, vol. 16, no. 2, pp. 101–114, 2002.

[22] Sietse M. Rispens, Mirjam Pijnappels, Kimberley S. van Schooten, Peter J. Beek, Andreas Daffertshofer, and Jaap H. van Dieën, "Consistency of gait characteristics as determined from acceleration data collected at different trunk locations," *Gait and Posture*, vol. 40, no. 1, pp. 187–192, 2014.

[23] "Estimation of gait cycle characteristics by trunk accelerometry," *Journal of Biomechanics*, vol. 37, no. 1, pp. 121–126, 2004.

[24] Mahesh Pal and Giles M. Foody, "Feature selection for classification of hyperspectral data by SVM," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 5, pp. 2297–2307, 2010.