

Automated tremor detection in Parkinson's Disease using accelerometer signals

Ada Zhang
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA
ajzhang@cs.cmu.edu

Rubén San-Segundo
Center for Information Processing
and Telecommunications
Universidad Politécnica de Madrid
Madrid, Spain
ruben.sansegundo@upm.es

Stanislav Panev
Griffin Tabor
Katelyn Stebbins
Andrew S. Whitford
Fernando De la Torre
Jessica K. Hodgins
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA

ABSTRACT

Wearable sensor technology has the potential to transform the treatment of Parkinson's Disease (PD) by providing objective analysis about the frequency and severity of symptoms in everyday life. However, many challenges remain to developing a system that can robustly distinguish PD motor symptoms from normal motion. Stronger feature sets may help to improve the detection accuracy of such a system. In this work, we explore several feature sets compared across two classification algorithms for PD tremor detection. We find that features automatically learned by a Convolutional Neural Network (CNN) lead to the best performance, although our handcrafted features are close behind. We also find that CNNs benefit from training on data decomposed into tremor and activity spectra as opposed to raw data.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning**: *Neural networks*; • **Human-centered computing** → *Ubiquitous and mobile computing*; • **Applied computing** → *Health informatics*;

KEYWORDS

Parkinson's Disease; wearable sensors; mobile health; deep learning

ACM Reference Format:

Ada Zhang, Rubén San-Segundo, Stanislav Panev, Griffin Tabor, Katelyn Stebbins, Andrew S. Whitford, Fernando De la Torre, and Jessica K. Hodgins. 2018. Automated tremor detection in Parkinson's Disease using accelerometer signals. In *ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE '18), September 26–28, 2018, Washington, DC, USA*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3278576.3278582>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHASE '18, September 26–28, 2018, Washington, DC, USA

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5958-0/18/09...\$15.00

<https://doi.org/10.1145/3278576.3278582>

1 INTRODUCTION

Approximately one million individuals in the United States are living with Parkinson's Disease (PD), a progressive neurological condition that most notably manifests in a variety of motor symptoms, such as tremor, slowness, or stiffness. Medication can effectively manage symptoms and improve quality of life. However, precise dosing is important: too small of a dose will fail to have a therapeutic effect, but too large of a dose can cause adverse side effects. The current standard of care calls for a specialist to assess PD patients every few months and adjust their medication based on self-reported symptoms and a short motor function evaluation. However, these assessments can be subjective and a short office visit may not accurately reflect a patient's experience at home. With the advent of inexpensive and widely-available mobile and wearable sensor technology, there has been increasing interest in developing systems for continuous monitoring of PD symptoms. Statistical and machine learning algorithms have recently shown promise in processing raw sensor data into disease-relevant information that is easily interpretable by physicians and patients. For example, on data collected from PD patients in a simulated home environment, Pulliam *et al.* [3] report reasonable agreement between sensor-derived estimates of tremor, bradykinesia, and dyskinesia severity and assessments from clinical experts.

In natural living conditions, however, it is particularly difficult for a system to accurately distinguish PD symptoms (such as tremor) from other daily activities (such as brushing one's teeth). One aspect of such a system is the choice of features extracted to represent the sensor signal. The majority of researchers in automated PD symptom detection have used variations on a standard set of features that are common in the field of activity recognition. In this work, we compare this standard set of features to four unexplored feature sets. All feature sets were compared across two different classification algorithms. In this way, we could analyze whether performance gains were due to improved features or stronger algorithms.

2 METHODS

Wrist-worn accelerometers were used to record movement data from six individuals with PD as they engaged in tasks that were chosen to mimic activities of daily living and clinical diagnostic tests. Videos of the data collection sessions were used to assign

Table 1: Area Under the Curve (AUC) and False Positive Rate (FPR) at 0.9 True Positive Rate (TPR) values for the five feature sets and two classifiers: Random Forest (RF) and Multi-layer Perceptron (MLP)

Feature Sets	AUC		FPR at 0.9 TPR	
	RF	MLP	RF	MLP
Baseline	0.830	0.829	0.45	0.45
MFCCs	0.851	0.853	0.40	0.39
CNN	0.850	0.850	0.38	0.41
MFCCs-T/A	0.869	0.870	0.33	0.33
CNN-T/A	0.884	0.887	0.32	0.30

labels of “tremor” or “no tremor” to every 3-second segment of accelerometer data.

We compared five distinct sets of features derived from the data:

- *Baseline* – This set consists of 90 features per axis that are common in previous works and include those used by Fisher *et al.* [2].
- *Mel frequency cepstral coefficients (MFCCs)* – These features are adapted from the speech processing field and have previously been shown to be effective for human motion understanding of accelerometer signals [4, 5]. Here, we evaluate their effectiveness on PD tremor detection.
- *MFCCs on tremor/activity spectra (MFCCs-T/A)* – In an attempt to enhance the tremor signal measured during movement, we developed a method to decompose the raw signal into separate tremor and voluntary movement (activity) spectra. This separating method was inspired by the main voice extraction technique proposed by Durrieu *et al.* [1], which uses non-negative matrix factorization. This feature set includes MFCCs computed on the raw signal, the tremor spectrum, and the activity spectrum.
- *CNN* – These features are taken from the last layer of a CNN trained on the raw signal, before the fully connected layers of the multilayer perceptron (MLP) classifier.
- *CNN on tremor/activity spectra (CNN-T/A)* – These features are taken from the last layer of a CNN trained on the separated tremor and activity spectra.

Feature sets were evaluated by comparing their performance using two different classifiers. For our baseline algorithm, we used a random forest (RF) because preliminary tests indicated that an RF performed at least as well as or better than other traditional machine learning algorithms, such as decision trees, 5-nearest neighbor, and Support Vector Machines. Our other classifier was an MLP, which is typically used with deep learning pipelines. Classifier and feature set performance was compared by computing the area under the curve (AUC) of the receiver operating characteristic (ROC) curves and the False Positive Rate (FPR) at a True Positive Rate (FPR) of 0.9.

3 RESULTS AND CONCLUSION

The most effective system for tremor detection from accelerometer signals was the full deep learning architecture – a CNN to learn features with an MLP at the end for classification – trained on the separated tremor and activity spectra (CNN-T/A with MLP). Table 1 compares the performance of several feature sets and classifiers. The difference in AUC values between the best- and worst-performing feature set/classifier pair was 0.058. The differences are more pronounced when comparing the FPR of the algorithms at 0.9 TPR, where the best system demonstrated a 0.15 reduction in FPR.

It is interesting to note that the main performance advantage of the deep learning architecture was through the learned features, and that the standard MLP classifier at the top of the network offered little to no performance advantage over other classifiers. In fact, the differences between RF and MLP classifiers given the same feature set (mean Δ : 0.002 in AUC and mean Δ : 0.0075 in FPR at 0.9 TPR) are an order of magnitude smaller than the difference in performance between feature sets given the same classifier (up to $\Delta = 0.054$ in AUC and $\Delta = 0.15$ reduction in FPR at 0.9 TPR).

Another observation is that, while the CNN-T/A features were able to offer a significant improvement over baseline features, the improvement was considerably smaller over our handcrafted MFCCs-T/A features. We suspect that, because the CNN-T/A features are learned from data, they may fail to generalize if the training dataset is small. Thus, for very small datasets, our handcrafted MFCCs-T/A features might outperform those learned by a CNN.

Finally, results indicate the importance of preprocessing the data to separate it into the tremor and activity spectra before inputting it into the CNN architecture. In fact, when a CNN is trained on raw data, performance is worse than our handcrafted MFCCs-T/A features.

In future work, we plan to analyze how sensitive these deep learning architectures are to the size of the training set, and to determine whether our handcrafted MFCCs-T/A features may offer an advantage over CNN-learned features on particularly small datasets. We also plan to analyze how well these features perform on data collected in more naturalistic environments.

REFERENCES

- [1] Jean-Louis Durrieu, Bertrand David, and Gaël Richard. 2011. A musically motivated mid-level representation for pitch estimation and musical audio source separation. *IEEE Journal of Selected Topics in Signal Processing* 5, 6 (2011), 1180–1191.
- [2] James M Fisher, Nils Y Hammerla, Thomas Ploetz, Peter Andras, Lynn Rochester, and Richard W Walker. 2016. Unsupervised home monitoring of Parkinson’s disease motor symptoms using body-worn accelerometers. *Parkinsonism & Related Disorders* 33 (2016), 44–50.
- [3] Christopher L. Pulliam, Dustin A. Heldman, Elizabeth B. Brokaw, Thomas O. Mera, Zoltan K. Mari, and Michelle A. Burack. 2018. Continuous Assessment of Levodopa Response in Parkinson’s Disease Using Wearable Motion Sensors. *IEEE Transactions on Biomedical Engineering* 65, 1 (2018), 159–164.
- [4] Rubén San-Segundo, Juan Manuel Montero, Roberto Barra-Chicote, Fernando Fernández, and José Manuel Pardo. 2016. Feature extraction from smartphone inertial signals for human activity segmentation. *Signal Processing* 120 (2016), 359–372.
- [5] Sebastián R Vanrell, Diego H Milone, and H Leonardo Rufiner. 2018. Assessment of homomorphic analysis for human activity recognition from acceleration signals. *IEEE Journal of Biomedical and Health Informatics* 22, 4 (2018), 1001–1010.