One Hip Wonder: 1D-CNNs Reduce Sensor Requirements for Everyday Gait Analysis

Jens Seemann[1 [0000−0001−7317−8967]], Tim Loris[1 [0009−0000−4810−6355]], Lukas Weber[1 [0009−0009−3172−8313]], Matthis Synofzik[2,3 [0000−0002−2280−7273]], Martin A. Giese[1 [0000−0003−1178−2768]], and Winfried Ilg[1,3 [0000−0002−1537−1336]]

[1] Section Computational Sensomotorics, Hertie Institute for Clinical Brain Research, Tübingen, Germany

[2] Division of Translational Genomics of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, Tübingen, Germany

[3] German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany

Abstract. Wearable inertial measurement units (IMU) enable large-scale multicenter studies of everyday gait analysis in patients with rare neurodegenerative diseases such as cerebellar ataxia. To date, the quantity of sensors used in such studies has involved a trade-off between data quality and clinical feasibility. Here, we apply machine learning techniques to potentially reduce the number of sensors required for real-life gait analysis from three sensors to a single sensor on the hip. We trained 1D-CNNs on constrained walking data from individuals with cerebellar ataxia and healthy controls to generate synthetic foot data and predict gait features from a single sensor and tested them in free walking conditions, including the everyday life of unseen subjects. We compare 14 stride-based gait features (e.g. stride length) with three sensors (two on the feet and one on the hip) with our approach estimating the same features based on raw IMU-data from a single sensor placed on the hip. Leveraging layer-wise relevance propagation (LRP) and transfer learning, we determine driving elements of the input signals to predict individuals' gait features. Our approach achieved a relative error ($< 5\%$) similar to the state of the art three-sensor approach. Thus, machine learning-assisted one-sensor systems can reduce the complexity and cost of gait analysis in upcoming clinical studies while maintaining clinical meaningful effect sizes.

Keywords: Clinical gait analysis · time-series · CNN · explainability · cerebellar ataxia

1 Introduction

In a wide variety of movement disorders, gait emerges as a cardinal symptom, which is often caused by progressive neurodegeneration. Accurate gait analysis in real life is crucial for the evaluation of upcoming treatments in such disorders, extracting disease-specific gait features like highly variable stride lengths and increased upper body sway [3]. Gait analysis conducted via multiple IMU-sensors on different parts of the body accurately extracts sensitive gait features but lacks clinical feasibility as it is costly and inconvenient for patients to use on a daily basis. While one-sensor systems deliver reliable information for average values of gait speed or stride length, for measures of spatio-temporal variability they have been demonstrated to be less reliable and less sensitive than their three-sensor counterparts, including an additional sensor on each foot [5].

1.1 Motivation

Hereditary cerebellar ataxia is a neurodegenerative movement disorder that causes progressive difficulty with walking and balance, resulting in decreased quality of life [14]. Spatio-temporal gait features extracted from wearable IMU-sensors indicate higher effect sizes for disease severity of ataxic patients in real-life in comparison to constrained walking trials [12], [23]. In particular, stride length variability has proved to be a strong indicator of whether and how fast the disease is progressing [11]. Gait features obtained with wearable IMU-sensors are thus promising response markers and can be used to validate future therapeutic approaches, as well as to enable online therapies [11]. Currently, no treatment for cerebellar ataxia exists, but gene therapies tested in mice are yielding promising results [6]. As cerebellar ataxia is a rare disease, validation of potential treatments often necessitates multicenter studies to ensure a sufficiently large number of subjects [22]. Reliable and feasible one-sensor systems are therefore particularly promising for upcoming therapy studies with real-life assessments, as these optimize patient convenience and reduce costs.

1.2 Objective

We hypothesize that machine learning can be used to predict various gait features represented only in hip data with high accuracy. Previous IMU-based machine learning approaches are limited to classification problems [13], are often based on constrained walking scenarios, [10] or clinically unfeasible sensor set-ups with 17 sensors [15]. Previously, 1D-CNNs have been used to predict five parameters simultaneously, based on a constrained walking dataset consisting of 1185 steps from 99 subjects, 54% of whom were diagnosed with a gait disorder or tendency to fall [8]. By contrast, we present an approach using 1D-CNN machine learning to reduce the number of sensors required for real-life gait analysis in cerebellar ataxia to one hip sensor only.

The aim of this study is to predict ataxic-sensitive gait features with small relative errors ($< 5\%$) using 1D-CNN-assisted one-sensor systems in everyday life. For an application in a clinical context it is crucial to find traceable mechanisms in gait feature prediction, which are revealed by explainability methods in the three- and one-sensor systems.
In this work, three approaches will be compared to investigate the impact of foot data on various stride-based gait features in subjects' everyday life:
(i) Using the complete set of three sensors (3S) resulting in 27 channels of IMU-data. (ii) Using the hip data and synthetic foot data (1S+2$\hat{S}$) resulting in 9 channels of orignal IMU-data + 18 channels of generated IMU-data. (iii) Only using the hip sensor (1S), resulting in 9 input channels of IMU-data. The different inputs are utilized in a 1D-CNN parameter prediction network (PN) of 14 stride-based gait features [2] (s. fig. 1): 7 spatio-temporal features (stride length, gait speed, stride duration, circumduction, lateral deviation, foot elevation and double support fraction), 4 foot angles (pitch at heel strike/midswing/toe off and toe out) and 3 upper body sway measures (range of motion in sagittal, transverse and coronal plane). In approach (ii) a second encoder-decoder based network serves as a foot data generation network (FN), aiming to generate foot sensor data from hip sensor data, which can be used as additional input for PN in a one-sensor scenario. Finally, we investigate LRP-activation patterns for regression models [17] to identify the relevant input channels for each gait feature.

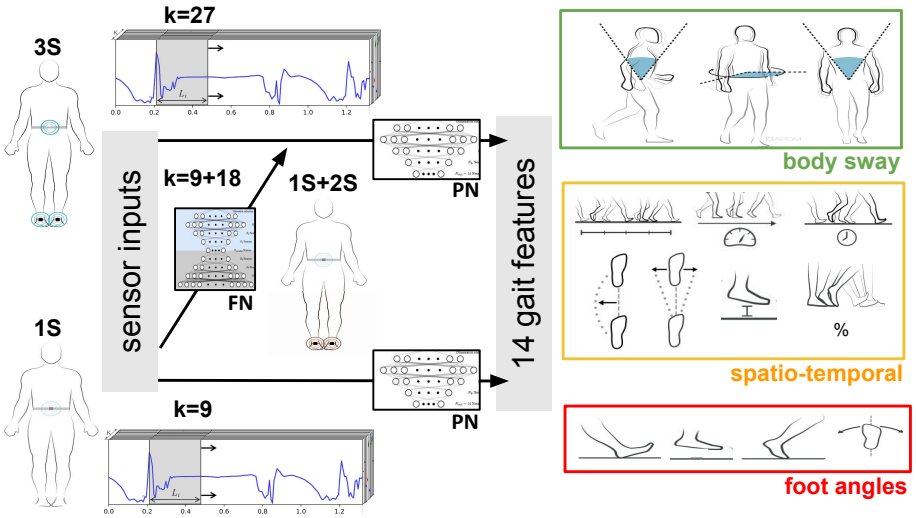


Fig. 1: Three approaches to predict 14 gait features from IMU-data using parameter prediction networks (PN) with different number of input channels k: Stride-based data from all three worn sensors (3S), one hip sensor (1S) and one hip sensor with synthetic foot data using a foot signal prediction network (FN) (1S+2$\hat{S}$). Gait features from: [2].

2 Methods

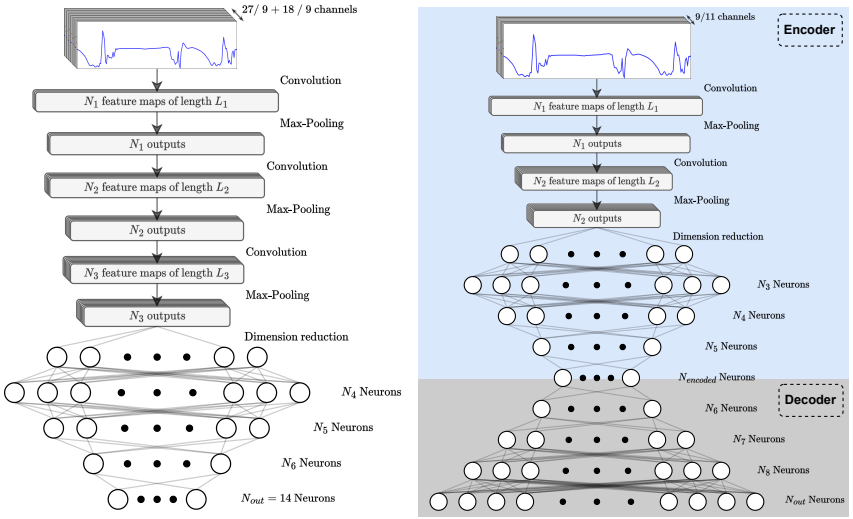
The dataset used here to train the two neural networks PN and FN is described in terms of data collection and clinically relevant details in [12] and [23].

2.1 Dataset

Gait data was collected from 47 healthy controls (HCs) and 59 patients with cerebellar ataxia (PATs) at a baseline, 1-year and eventually 2-years follow-up assessment by using three Opal sensors (APDM, v2), placed at the lumbar region and on both shoes. Each Opal sensor included triaxial accelerometers, gyroscopes and magnetometers with a fps = 128. Stride-based gait features were extracted from raw data using Mobility Lab (APDM, v2), validated in [19]. Each of the 106 subjects completed a constrained 50 m straight walk (NW). A subset of subjects (n = 87) participated in a 10 minute observed free walk (OW) in- and outside the clinics including stairs and busy hallways. A subset of OW-subjects (n = 64) completed several hours of unsupervised home recordings reflecting individual daily living (DW). All sets contained an equal proportion of HCs and PATs.
Preprocessing of the raw data included segmenting the gait sequences into separate strides based on one-sided heel strike events (HS), removing exceptionally short or long strides, and inserting the IMU-data into channels with a fixed size of 170 data points (1.33 seconds). If a stride was shorter, it was inserted in the middle of the channel and filled up with original sensor data of the previous and next stride, resulting in a total of 10,715, 61,877, and 427,519 strides per gait task, respectively.

2.2 Network architectures

Three PNs with different input channels k were trained to predict 14 gait features from different sensor-setups: using raw data from the hip sensor only (1S, k=9), the hip plus FN-generated feet data (1S+2$\hat{S}$, k=9+18) in comparison to the complete set of all three sensors (3S, k=27). Using 1D convolutions layers, the shifting of the kernels is done on the time dimension with a step size of one. The PN consisted of three convolutional layers, with a max-pooling layer with a size of 2 following each layer and three fully connected layers (s. fig. 2 (a)). The convolutional layers contained $N_1 = 32, N_2 = 64$ and $N_3 = 128$ kernels of size $L_1 = 30, L_2 = 15$ and $L_3 = 7$. The fully connected layer consisted of $N_4 = 4096, N_5 = 2048$ and $N_6 = 1024$ neurons and a corresponding number of bias values. The output layer of the network contained $N_{out} = 14$ neurons, corresponding to the 14 gait features. The FN was composed of an encoder and a decoder (s. fig. 2 (b)). The encoder consisted of three convolutional layers, each followed by a max-pooling layer with window size two and four dense layers, resulting in $N_{encoded} = 750$. The decoder contained four dense layers. The neuron number of the readout layer N_{out} was $170 \cdot 9 \cdot 2 = 3060$ neurons and a matching number of bias values.



(a) **PN** (b) **FN**

Fig. 2: Architectures of the 1D-CNN networks for (a) parameter prediction (PN) and (b) foot signal prediction (FN), modified from [8].

2.3 Training

In the supervised training phase of all three PNs, we used 14 stride based gait features previously determined by a validated three-sensors algorithm as ground truth [19]. We selected a training scheme simulating a clinical trial: A base network was trained on a larger set of constrained walking data, and is then re-trained with limited free walking data from new patients and tested on everyday data of the respective patients. Firstly, a 70:30 split of NW data set was done for training a base net for 1500 epochs. Next, the free walking data (OW) of unseen subjects was used as retraining data for 750 epochs. Finally, 1% of the DW data of the test subjects was utilized in a second retraining for 1500 epochs to explore further improvements.

For the PN network the learning rate α was $5 \cdot 10^{-6}$ with normal distribution ($\mu = 0,\ \sigma = 1$) as weight initialization. Adam optimization algorithm [16] was performed with $\beta_1 = 0.9$, $\beta_2 = 0.99$, and $\epsilon = 10^{-8}$ and a mini-batch size of 64. To prevent overfitting, the network was trained with a weight decay of 10^{-5}. To facilitate the training of the networks, all outputs were normalized between 0 and 1.

For the FN network the learning rate α was changed to $8 \cdot 10^{-7}$ with He initialization [9]. Additionally, gradient clipping was used with a value of 1 [21]. Soft-DTW from [4] and implementation of [18] was used as the loss function to quantify how similar the predicted foot data was to the original one [20].

2.4 Layerwise Relevance Propagation

For a deeper analysis of the networks of this work, LRP was utilized, which described on the basis of relevance scores what kind of influence the respective sensor input had on the prediction. LRP was implemented using the 'Epsilon-Plus' function of the zennit framework [1]. Thus, for convolutional layers the Alpha-Beta rule with $\alpha = 1,\ \beta = 0$, and for fully connected layer the Epsilon rule with $\epsilon = 1 * 10^{-6}$ was applied. To allow a specific interpretation of the inputs, the PNs were trained from scratch on only one gait features instead of all 14 for 500 epochs and were retrained for 150 epochs.

3 Results

In all approaches (1S), (1S+2$\hat{S}$), and (3S), we investigate whether transfer learning from an existing dataset of restricted walking trials (NW) to new patients based on a small set of recorded gait samples with 3 sensors in the clinics (OW) is beneficial to generalize to unseen real-life data (DW). First, the quality of the generated foot data is shown, then the predictions of the different gait features are compared, and finally a revealed mechanism for stride length is described as an example for LRP in gait analysis.

3.1 Generation of foot trajectories

Figure 3 visualizes the predictions of the networks that were trained on NW or OW data exclusively and those that were trained on NW and retrained with OW data. All networks were tested on the DW dataset. The foot data channels are estimated correctly on average, but the variance shows clear deviations: All trained networks overestimate the data of patients (s. fig. 3, top). Especially, the variance in the predicted x-axis of the accelerometer indicates huge noise, that is not present in the original data. This is primarily during stance phase (0.3 s - 0.75 s) where one foot is on the ground and thus not accelerating at all. Furthermore, the noise of the accelerometer is lower with HC compared to the patient group. Retraining on OW lower the standard deviations of the predicted DW curves and better fit the distribution of the data for HCs and PATs. The network trained on the OW data only shows the strongest fluctuations of all three networks for PATs. Since the network trained with the NW dataset and retrained with the OW dataset has the lowest error and makes the best predictions of the three networks on everyday data, this is further used to predict everyday foot data for the 1S+2$\hat{S}$ approach. The network trained on the OW data only shows the strongest fluctuations of all three networks for PATs. Since the network trained with the NW dataset and retrained with the OW dataset has the lowest error and makes the best predictions of the three networks on everyday data, this is further used to predict everyday foot data for the 1S+2$\hat{S}$ approach.

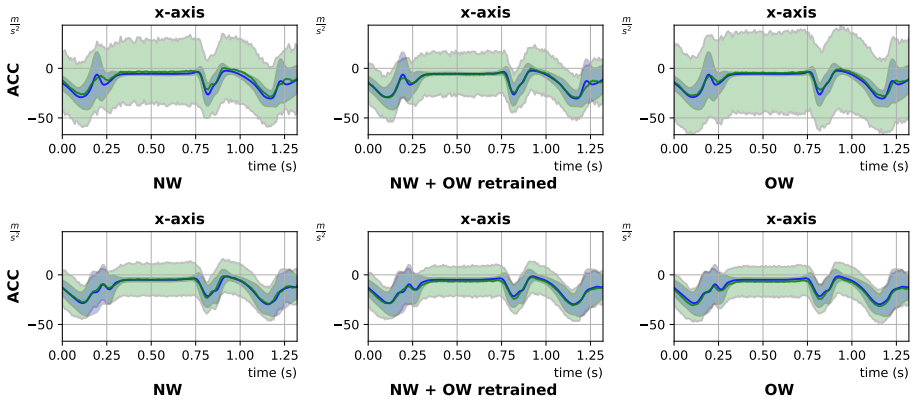


Fig. 3: Mean and standard deviation of original DW foot data (blue) and FN predictions (green) trained on NW data, and retrained on OW, and trained on OW exclusively for patients (top) and healthy controls (bottom).

3.2 Prediction of gait features

First results deliver low (<5%) relative errors for stride duration, gait speed and stride length, using hip raw data only after retraining with OW-data (s. tab. 1). For these gait features, using the complete set of sensor data shows similar results (e.g. stride duration rel. err: 0.72% (3S), 1.19% (1S)). All gait features benefit from an additional retraining on 1% of everyday data (e.g. stride length rel. err: 4.37% (1S, OW), 2.35% (1S, DW1%)).
The foot angles at TO and HS show small relative errors of 2.47% and 5.17% in the three-sensor approach, but can also be captured with medium deviations by the one-sensor approach (6.61% and 10.22% resp.). The relative error decreases to 3.16% resp. 6.90% when an additional 1% of the everyday data is used for retraining. For other foot related gait features like lateral deviation or toe-out angle, an accurate estimation seems to be difficult with all our approaches.
Predicting sway of the upper body reveals moderate relative errors (∼10%) in both conditions (e.g. coronal sway rel. err.: 8.81% (1S), 11.1% (3S), transverse sway rel. err.: 10.02% (1S), 8.96% (3S)).
The estimation of almost all gait features is improved by generating foot trajectories from hip trajectories. Here, for example, the rel. error of the stride length decreases from 4.37% (1S) to 4.05%. Thus, using synthetic foot data can achieve mild improvements for gait feature predictions.
In terms of mean error and standard deviation per subject, it can be seen that for the stride length predominantly the outliers of the base nets lie closer to the straight lines after retraining and thus correspond better to the ground truth (s. fig. 4). In general, subjects-wise stride length standard deviations lie mostly below the ground truth, so that the nets show a lower standard deviation per subject in their estimates.

Table 1: Relative errors of gait feature predictions for the sensor set-ups after retraining on OW and 1% of DW data set and tested on DW data set.

Rel. error	real feet (3S)		pred. feet(1S+2$\hat{S}$)		hip(1S)	
after retraining	OW	DW1%	OW	DW1%	OW	DW1%
stride duration [s]	0.72%	0.49%	1.39%	0.95%	1.19%	0.93%
gait speed [m/s]	2.53%	1.41%	3.82%	2.25%	4.13%	2.36%
stride length [m]	2.56%	1.49%	4.05%	2.18%	4.37%	2.35%
double support	3.26%	2.18%	7.70%	4.14%	7.74%	4.15 %
circumduction [m]	17.31%	14.24%	35.81%	28.70%	38.23%	28.77%
foot elevation [m]	32.00%	22.83%	39.57%	30.20%	41.88%	31.91%
lateral dev. [m]	49.82%	37.19%	74.85%	64.39%	72.47%	62.49%
pitch at TO [°]	2.47%	1.52%	6.25%	3.16%	6.61%	3.16 %
pitch at HS [°]	5.17%	3.50%	10.64%	7.00%	10.22%	6.90%
pitch at MS [°]	10.57%	8.17%	21.87%	15.53%	22.98%	15.18%
toe out[°]	34.77%	23.16%	84.41%	52.43%	85.43%	52.03%
transverse ROM [°]	8.96%	6.82 %	11.35%	7.61%	10.20%	7.11%
coronal ROM [°]	11.11%	6.73 %	9.88%	5.70%	8.81%	5.80%
sagittal ROM [°]	26.31%	19.21%	26.46%	18.06%	26.05%	18.86%

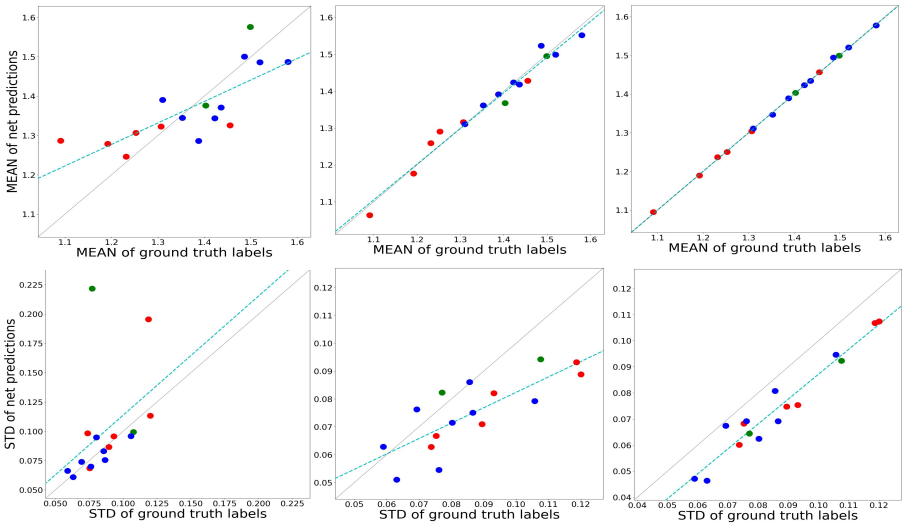


Fig. 4: Means (top) and standard deviations (bottom) of the 1S network and corresponding ground truth labels per subject for the stride length, comparing the base (left), OW-retrained (middle) and DW1%-retrained (right) networks. Blue: HCs, Red: PATs.

However, foot-specific gait parameters not in gait direction such as lateral deviation, circumduction, toe out and foot elevation are inaccurately predicted with large relative errors in the single sensor condition (> 30%), but also while using all of the three sensors (> 15%) (s. tab. 1).

3.3 Explainability

For stride length prediction in 3S, strongly positively relevant areas of the x and z axis of the accelerometer are specifically located in the swing phase (starting at 60% of the gait cycle) of the referenced foot (s. fig. 5, top). The gyroscope does not appear to be used in any of the three sensor locations for stride length prediction. The magnetometer of the hip in x and z direction shows slight relevances during the whole swing phase.
If only the hip sensor (1S) is available, the positive relevances scatter in the whole range of the gait cycle, mainly occurring in the x and z channel of the accelerometer and magnetometer (s. fig. 5, bottom). The importance of the magnetometer is overall higher compared to 3S. The gyroscope is not relevant for stride length prediction in the 1S approach.

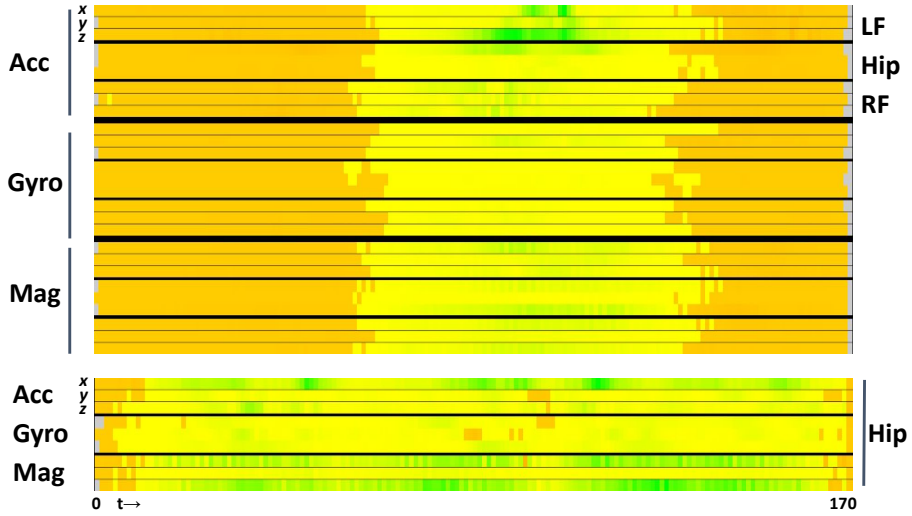


Fig. 5: Example stride with high (green), low (yellow) and slightly negative (orange) relevance scores predicting stride length with true label 1.31 m, and predicted stride length 1.32 m (3S, top) and 1.38 (1S, bottom), respectively. Rows: LRP extracted relevance scrores for IMU-sensor inputs with a length of 170 of the triaxial accelerometers (Acc), gyroscopes (Gyro) and magnetometers (Mag) of the OPAL-sensors placed on the left foot (LF), hip and right foot (RF).

4 Discussion

Our key finding is that 1D-CNN based one-sensor systems are able to predict treatment responsive gait features with low relative error ($< 5\%$) in everyday life. Furthermore, LRP indicates that gait features are traceably represented in the hip when foot sensors are not available, facilitating assessment by clinicians. For the analysis of everyday data, retraining with data previously recorded in the clinics is beneficial: key gait features responsive to ataxia, such as stride length, are predicted accurately with just one sensor on the hip. Since cerebellar ataxia is characterized by increased variability of gaits features, the variance of predicted values is highly relevant. Using retraining on individuals, both PN and FN networks are able to recognize individual peculiarities in real-life gait. This might be beneficial as different patients show individual gait patterns or compensation strategies. Additional retraining with a small portion of the everyday data was shown to be further beneficial. However, since this would require wearing three sensors to generate the ground truth in everyday life, this approach is not applicable in the clinics.

Unsurprisingly, accurate prediction of some foot-related features like the toe-out angle requires feet sensors due to insufficient representation of the degrees of freedom in the hip, knee and ankle joints, in the single hip sensor. Gait features such as lateral deviation that relate to changes between two strides are inadequately represented by single stride data. Implementing LSTMs or transformer models, future approaches should incorporate longer inputs with additional contextual information to improve the quality of predictions for lateral deviation.

Before gait parameters are estimated in 1S+2$\hat{S}$, foot data is generated first, which show qualitative differences from the raw data. Before retraining the foot sensor predictions include significant noise on the x-axis of the ACC, in particular in the patient group. A possible reason for this could be highly altered gait patterns in the patient group with varying severities of ataxia. Improving the FN network can lead to using already existing three sensor based analysis pipelines with original hip data and synthetic foot data.

Analysing the 3S approach, LRP reveals acceleration in the direction of gait as highly relevant, which is initially read on the z axis due to the plantar flexion during early swing phase and primarily on the x axis during mid swing. This is an intuitively correct approach, because the stride length is defined as the distance covered by the referenced foot during the swing phase. The learnt mechanism of the 1S network to predict the stride length from the hip acceleration in vertical direction is consistent with conventional approaches using inverted pendulum models [7]. Biomechanically, larger strides tend to cause a larger displacement of the hip in the vertical direction during the gait cycle. In this respect, traceable areas of input are of great relevance in both the 1S and 3S approach.

For upcoming clinical trials, predicted gait features should be consistent with effect sizes from validated three sensor systems and surpass state of the art methods based on one sensor only, allowing everyday measurement of a larger number of patients for future therapy studies.

5 Conclusion

1D-CNN-assisted one-sensor systems are able to predict gait features from one hip raw data with low relative errors. For specific gait features, they do exhibit expected qualitative deficiencies compared to the use of three sensors, but these can be reduced by retraining on the individual patients. Explainability methods like LRP reveal meaningful mechanisms behind three- and one-sensor based gait feature predictions for clinicians and therapists. This work shows how established three-sensor based gait features from the laboratory can be transferred to everyday life by one sensor only, maximizing the clinical feasibility for future clinical studies.

Acknowledgments

The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Jens Seemann. This work was supported by Else Kröner-Fresenius-Stiftung: Project ClinbrAIn. Further support was received by the European Research Council ERC 2019-SYG under EU Horizon 2020 research and innovation programme (grant agreement No. 856495, RELEVANCE). ChatGPT generated the part of the title 'One Hip Wonder' given the abstract and the prompt to generate a fun title.

References

1. Anders, C.J., Neumann, D., Samek, W., Müller, K.R., Lapuschkin, S.: Software for dataset-wide xai: From local explanations to global insights with Zennit, CoRelAy, and ViRelAy. CoRR **abs/2106.13200** (2021)
2. APDM: Mobility lab whitepaper (2015), https://apdm.wpengine.com/wp-content/uploads/2015/05/02-Mobility-Lab-Whitepaper.pdf
3. Buckley, E., Mazzà, C., McNeill, A.: A systematic review of the gait characteristics associated with cerebellar ataxia. In Gait Posture **60**, 154–163 (2018)
4. Cuturi, M., Blondel, M.: Soft-dtw: a differentiable loss function for time-series. In: International Conference on Machine Learning (2017)
5. Czech, M., Demanuele, C., Erb, M.K., Ramos, V., Zhang, H., Ho, B., Patel, S.: The impact of reducing the number of wearable devices on measuring gait in parkinson disease: Noninterventional exploratory study. JMIR Rehabil Assist Technol **7**(2), e17986 (Oct 2020)
6. Ghanekar, S.D., Kuo, S.H., Staffetti, J.S., Zesiewicz, T.A.: Current and emerging treatment modalities for spinocerebellar ataxias. Expert Review of Neurotherapeutics **22**(2), 101–114 (2022), pMID: 35081319
7. Goyal, P., Ribeiro, V.J., Saran, H., Kumar, A.: Strap-down pedestrian dead-reckoning system. In: International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2011. pp. 1–7. IEEE / Institute of Electrical and Electronics Engineers Incorporated (2011)
8. Hannink, J., Kautz, T., Pasluosta, C.F., Gasmann, K.G., Klucken, J., Eskofier, B.M.: Sensor-based gait parameter extraction with deep convolutional neural networks. IEEE Journal of Biomedical and Health Informatics **21**(1), 85–93 (jan 2017)

9. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: 2015 IEEE International Conference on Computer Vision (ICCV). pp. 1026–1034 (2015)
10. Hossain, M.S.B., Dranetz, J., Choi, H., Guo, Z.: Deepbbwae-net: A cnn-rnn based deep superlearner for estimating lower extremity sagittal plane joint kinematics using shoe-mounted imu sensors in daily living. IEEE journal of biomedical and health informatics **26**(8), 3906–3917 (2022)
11. Ilg, W., Müller, B., Faber, J., van Gaalen, J., Hengel, H., Vogt, I.R., Hennes, G., van de Warrenburg, B., Klockgether, T., Schöls, L., Synofzik, M., the ESMI Consortium: Digital gait biomarkers allow to capture 1-year longitudinal change in spinocerebellar ataxia type 3. In Movement Disorders **37**(11), 2295–2301 (2022)
12. Ilg, W., Seemann, J., Giese, M., Traschütz, A., Schöls, L., Timmann, D., Synofzik, M.: Real-life gait assessment in degenerative cerebellar ataxia: Toward ecologically valid biomarkers. In Neurology **95**(9), e1199–e1210 (2020)
13. Jabri, S., Carender, W., Wiens, J., Sienko, K.H.: Automatic ml-based vestibular gait classification: examining the effects of imu placement and gait task selection. Journal of NeuroEngineering and Rehabilitation **19**(1), 132 (2022)
14. Joyce, M.R., Nadkarni, P.A., Kronemer, S.I., Margron, M.J., Slapik, M.B., Morgan, O.P., Rosenthal, L.S., Onyike, C.U., Marvel, C.L.: Quality of life changes following the onset of cerebellar ataxia: Symptoms and concerns self-reported by ataxia patients and informants. Cerebellum (London, England) **21**(4), 592–605 (2022)
15. Kadirvelu, B., Gavriel, C., Nageshwaran, S., Chan, J.P.K., Nethisinghe, S., Athanasopoulos, S., Ricotti, V., Voit, T., Giunti, P., Festenstein, R., Faisal, A.A.: A wearable motion capture suit and machine learning predict disease progression in friedreich's ataxia. Nature medicine **29**(1), 86–94 (2023)
16. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. CoRR **abs/1412.6980** (2014)
17. Letzgus, S., Wagner, P., Lederer, J., Samek, W., Müller, K.R., Montavon, G.: Toward explainable ai for regression models. IEEE Signal Processing Magazine **39**(4), 40–58 (2022), http://arxiv.org/pdf/2112.11407v2
18. Maghoumi, M., Taranta, E.M., LaViola, J.: Deepnag: Deep non-adversarial gesture generation. In: 26th International Conference on Intelligent User Interfaces. pp. 213–223 (2021)
19. Morris, R., Stuart, S., McBarron, G., Fino, P.C., Mancini, M., Curtze, C.: Validity of mobility lab (version 2) for gait assessment in young adults, older adults and parkinson's disease. Physiological measurement **40**(9), 095003 (2019)
20. Müller, M.: Dynamic Time Warping, pp. 69–84. Springer Berlin Heidelberg, Berlin, Heidelberg (2007)
21. Pascanu, R., Mikolov, T., Bengio, Y.: On the difficulty of training recurrent neural networks (2012)
22. Ruano, L., Melo, C., Silva, M.C., Coutinho, P.: The global epidemiology of hereditary ataxia and spastic paraplegia: a systematic review of prevalence studies. Neuroepidemiology **42**(3), 174–183 (2014)
23. Thierfelder, A., Seemann, J., John, N., Harmuth, F., Giese, M., Schüle, R., Schöls, L., Timmann, D., Synofzik, M., Ilg, W.: Real-life turning movements capture subtle longitudinal and preataxic changes in cerebellar ataxia. Movement disorders : official journal of the Movement Disorder Society **37**(5), 1047–1058 (2022)