Supplementary Material for: Gait Motion Classification for Neurodegenerative Diseases by Recurrence Structure Analysis

Gauthier Debes, Diwei Wang, Peter beim Graben, Frédéric Blanc, Candice Muller, Hyewon Seo and Axel Hutt

I. Complexity measures

A. Conventional measures

Symbolic sequences exhibit a certain complexity or regularity and multiple complexity measures are already known (1; 2; 3). For instance, the *alphabet cardinality* (4; 3) $M(\varepsilon)$counts the number of distinct states that appear in the sequence. For illustration, let us assume the symbolic sequence

$$[0,0,1,1,1,0,2,2,0,0,3,0,0,0,1,1,1,0,2,2,\\ 0,0,3,0,0,0,1,0,2,2,0,4,4,4]\ . \quad (1)$$

There are five distinct states $0,1,2,3,4$ and thus alphabet size $M=5$.

Moreover, the *number of words* may also express the degree of complexity. It represents the number of distinct states with different size in the sequence. Accordingly, the sequence (1) has the distinct states $0,00,000,1,111,22,3,444$, thus $C_{\text{number of words}}=8$.

The Lempel-Ziv Complexity, $C_{\text{Lempel}-\text{Ziv}}$ (5; 6) ,evaluates sub-sequences in the global sequence and counts how many distinct ones exist. For illustration, we modify the symbolic sequence (1) to include only the sequence of state appearances. The new sequence reads

$$[0,1,0,2,0,3,0,1,0,2,0,3,0,1,0,2,0,4] \quad (2)$$

yielding the Lempel-Ziv sub-sequences $[0],[1],[0,2],[0,3],[0,1,0,2,0,3,0],[1,0,2,0,4]$ and the corresponding $C_{\text{Lempel}-\text{Ziv}}=6$.

B. New measures

Now we introduce several different complexity measures, which serve as biomarkers for the gait sequence:

- As stated before, a regular gait exhibits a single metastable state in each step and we evaluate the degree of regularity as the difference between the number of metastable states' appearances n_k in segment k and the number of steps. Then, the average over all metastable states n yields $C_{\text{appearence}}=\sum_{k=1}^{N}|n_k-3|\,/3n$ .For instance, for the sequence (2) $C_{\text{appearance}}=(|3-3|/3+|3-3|/3+|2-3|/3+|1-3|/3)/4=0.25$.
- The previous complexity measure estimates the deviation of the optimal number of metastable states. The symbolic sequence of a perfectly regular gait segmented in three steps is supposed to contain three times the same states' pattern. Now we quantify the states' pattern and count the number of different sub-sequences with only distinct metastable states. For a perfectly regular gait, the sequence exhibits three times a unique sub-sequence and the corresponding complexity measure should be equal to 1. The resulting number of sub-sequences $C_1=\#(\text{Sub}-\text{sequences})$ is an additional complexity measure. New sub-sequences start with a metastable state only and transient states at the beginning are neglected. In addition, the first state at the end of a sequence is neglected. For illustration, the sequence

(2) yields the corresponding sub-sequences $[0,1,0,2,0,3,0]$, $[1,0,2,0,3,0]$, $[1,0,2,0,4]]$, where two sub-sequences are different. Hence the new complexity measure is $C_1 = 2$.

- Extending the previous complexity measure, one searches for the longest sub-sequence that occurs more than once in the symbolic sequence. Once identified, this sub-sequence is removed from the sequence. Then, the algorithm aims to find the second longest recurrent sub-sequence, and continues in this manner. When only non-recurrent sub-sequences remain, the algorithm identifies the longest sub-sequence composed only of distinct metastable states, neglecting transient states. This iterative process continues until no symbols remain in the symbolic sequence. Similar to the algorithm of the previous complexity measure C_1, the first transient state and the last metastable state are removed, if necessary, yielding another complexity measure, $C_2 = \#(\text{Iterated Sub} - \text{sequences})$.Using sequence (2) as an example, the algorithm identifies the sequence $[1,0,2,0,3,0]$ as the longest sub-sequence occurring at least twice. Subsequently, only non-recurrent metastable states $[1,0,2,0,4]$ remain and new complexity measure is $C_2 = 2$.
- Metastable states may occur regularly indicating a certain regularity of the full sequence. We calculate the deviation of the sequence position distance between the occurrences of the states. To this end, we segment the symbolic sequence step by step and keep only those metastable states, which appear only once in each step. Considering the center sequence position of each state, computing the distance between the three center positions allows to compute the average deviation of the center positions. For illustration, the sequence (1) yields the sub-sequences $[0,0,1,1,1,0,2,2,0,0,3]$, $[0,0,0,1,1,1,0,2,2,0,0,3]$, $[0,0,0,1,0,2,2,0,4,4,4]$. Then, keeping the states '1' and '2' only because they appear once in each step, their center positions are $[4,16,27]$ for state '1' and $[7.5,19.5,29.5]$ for state '2'. The distances between them are $[12,11]$ for state '1' and $[12,10]$ for state '2' leading to the corresponding deviations 0.5 and 1 and the final occurrence deviation is $C_{\text{deviation}} = 0.75$.

For illustration, Table I shows examples of various complexity measures considered in the work.

	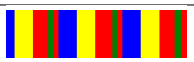	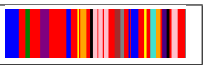
Alphabet Size	4	12
Lempel-Ziv Complexity	7	18
$C_{\text{appearance}}$	0	0.48
C_2	1	9
$C_{\text{deviation}}$	0.72	3.65
$C_{\text{step period}}$	48.0	91.4

Table I: **Example of complexity measures on two different symbolic sequences.** Left : Healthy controls. Right : Mild Dementia with Lewy Bodies. Complexity measures as follows: M is the alphabet cardinality, $C_{\text{Lempel}-\text{Ziv}}$ is the Lempel-Ziv complexity, $C_{\text{appearance}}$ is the appearance complexity, C_2 is the number of iterated sub-sequences, $C_{\text{deviation}}$ the deviation complexity and $C_{\text{step period}}$ is the step period complexity.

II. Information on data

Table II provides the number of datasets.

III. Network configurations

In general, one can assume a dataset composed of X subjects, and for each subject, one has N biomarkers that quantify the gait impairments. These N biomarkers are used as inputs for the classification. We use the *Tensorflow* library (7) to build a Multi-Layer Perceptron (8) (MLP) composed of four layers. At first, we standardize the X values of each biomarker. Mathematically, for each of the N biomarkers, the function is $z_i = (x_i - \mu)/\sigma,\ i = 1,\ldots,X$ with x_i the value to standardize, z_i the standardized value, μ the mean, and σ the standard deviation of the X values of

class	description	number of datasets
Gait score	Normal walk	65
	Slight impairment	148
	Mild impairment	55
	Moderate impairment	74
Diagnosis	Healthy	22
	mild DLB	63
	mild DLB	76
	mild AD	112
	severe AD	69
	total	342

Table II: **The number of datasets for each gait and diagnosis class.** The number of dataset is different from the number of video and thus subjects, since each video may include several time windows to analyze.

the biomarker n_i in the dataset. The values z_i are logits, which represent the scores of the classification model. This normalization ensures, that all inputs have zero mean and unity standard deviation, giving them equal importance in the classification. The first three layers of the neural network employ the $\mathrm{ReLU}(x) = \max(0, x)$ as activation function, while the final layer uses a softmax activation function to express classifications as probabilities(8) $\hat{y}_{ij} = e^{z_i} / \sum_{k=1}^{X} e^{z_k}$, with $i = 1, \ldots, X$ and $j = 1, \ldots, C$. Here, C is the number of outputs of layer 4 and hence the number of possible classes, e.g. $C = 2$ in the case of the classification *regular/irregular*.

We trained the four-layer MLP to minimize cross-entropy loss using gradient descent, utilizing the Adamax optimizer to adjust the learning rate based on the moments of the gradients calculated through backpropagation. Adamax, a variant of the Adam optimizer, is known for providing greater stability in the presence of noisy data. The cross-entropy loss is calculated by $L_{ce} = -1/N \sum_{i=1}^{X} \sum_{j=1}^{C} y_{ij} \log(\hat{y}_{ij})$, where X the number of data, C the number of classes, y_{ij} the real (target) probability for the data i to be in the class j (i.e. $y_{ij} = 1$ if the data is from the class j or $y_{ij} = 0$ otherwise.), and $\hat{y}_{ij}$ the predicted probability for this data to be in the class j.

Fig. 1 illustrates our four-layer MLP network after data normalization.

IV. Details on classifications

Table III provides p-values for the classification of various joints combinations and Table IV gives the p-values of surrogate tests for selected joints combinations. Finally, Table V provides an importance ranking of all complexity measures determined by a feature ablation test.

References

[1] C. Adami and N.J. Cerf. Physical complexity of symbolic sequences. Physica D, 137(1-2):62–69, 2000.

[2] F. Von Wegner, M. Wiemers, G. Hermann, I. Tödt, E. Tagliazucchi, and H. Laufs. Complexity measures for eeg microstate sequences: Concepts and algorithms. Brain Topogr., 37:296–311, 2024.

[3] P. beim Graben, K.K. Sellers, F. Fröhlich, and A. Hutt. Optimal estimation of recurrence structures from time series. Europhys. Lett., 114(3):38003, 2016.

[4] S. Janson, S. Lonardi, and W. Szpankowski. On average sequence complexity. Theoret. Comp. Sci., 326:213–227, 2004.

[5] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. IEEE Transactions on Information Theory, 22:75–81, 1976.

[6] A. Lempel and J. Ziv. On the complexity of finite sequences ieee transactions on information theory. IEEE Transactions on Information Theory, 23:337–343, 1977.

[7] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

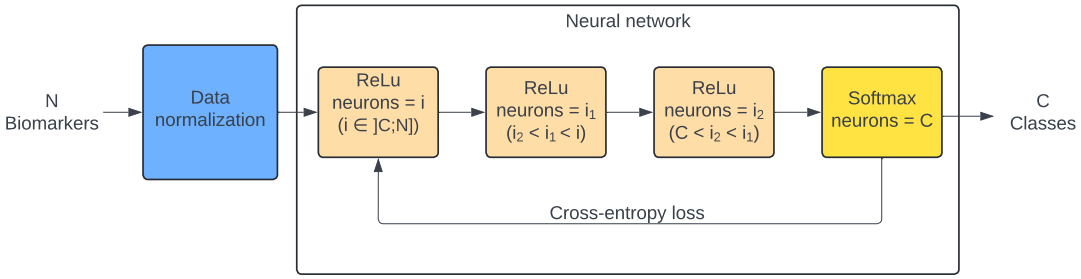

Figure 1: **Classification network configuration.** The network begins with N neurons, and each layer reduces the dimensionality of the input data. The first layer has N input neurons and i output neurons, the next layer has i input layers and i_2 output layers, the third layer has has i_2 input layers and C output layers. The last layer compute the softmax function. Parameters N and the number of classification classes C depend on the classification problem being solved and the dataset used (See section on classification in the main manuscript).

regular/irregular								
	4BPs	**4BPo**	**2Ls**	**2Lo**	**2As**	**2Ao**	**2Lo&2Ao**	**All**
4BPs		**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
4BPo	0.00		**0.00**	**0.00**	**0.00**	**0.01**	0.15	**0.00**
2Ls	**0.00**	**0.00**		0.83	**0.00**	**0.00**	**0.00**	**0.00**
2Lo	**0.00**	**0.00**	0.83		**0.00**	**0.00**	**0.00**	**0.00**
2As	**0.00**	**0.00**	**0.00**	**0.00**		0.63	**0.00**	0.21
2Ao	**0.00**	**0.01**	**0.00**	**0.00**	0.63		**0.00**	0.16
2Lo&2Ao	**0.00**	0.15	**0.00**	**0.00**	**0.00**	**0.00**		**0.00**
All	**0.00**	**0.00**	**0.00**	**0.00**	0.21	0.16	**0.00**	

AD/DLB								
	4BPs	**4BPo**	**2Ls**	**2Lo**	**2As**	**2Ao**	**2Lo&2Ao**	**All**
4BPs		**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
4BPo	**0.00**		**0.00**	**0.00**	**0.00**	0.93	**0.02**	**0.00**
2Ls	**0.00**	**0.00**		**0.03**	**0.00**	**0.00**	**0.00**	0.06
2Lo	**0.00**	**0.00**	**0.03**		**0.00**	**0.00**	**0.00**	0.45
2As	**0.00**	**0.00**	**0.00**	**0.00**		**0.00**	**0.01**	**0.00**
2Ao	**0.00**	0.93	**0.00**	**0.00**	**0.00**		**0.00**	**0.00**
2Lo&2Ao	**0.00**	**0.02**	**0.00**	**0.00**	**0.01**	**0.00**		**0.00**
All	**0.00**	**0.00**	0.06	0.45	**0.00**	**0.00**	**0.00**	

Table III: **P-values for the classification of various joints combinations.**Two-sided Mann-Whitney U test results for each joints combinations on both classifications criteria. The p-value of a cell $[i; j]$ represents the probability that accuracy distribution of set i is not different to the distribution j. Significantly different distributions are marked in bold font. The significance level is $\alpha = 0.05$.

[8] C.M. Bishop. Pattern Recognition and Machine Learning. New York: Springer. ISBN 978-0-387-31073-2., 17 August 2006.

	4BPs	**4BPo**	**2Ls**	**2Lo**	**2As**	**2Ao**	**2Lo&2Ao**	**All**
Regular/Irregular)	**0.00**	0.39	**0.04**	**0.00**	0.716	0.49	0.13	0.28
AD/DLB	**0.00**	**0.00**	**0.00**	0.13	**0.00**	**0.00**	**0.00**	0.08

Table IV: **P-values of surrogate tests for selected joints combinations on both classification criteria.** P-values < 0.05 (bold font) indicate that the distribution of original classification results is different to the distribution of classification results from surrogate data. The surrogate data represents the same dataset but with randomized gait score (Regular/Irregular) and diagnosis (AD/DLB) labels.

regular/irregular			
Importance Rank	Feature removed	ΔA (in %)	A (in %)
1	$C_{\text{step period}}$	2.44	85.58
2	$C_{\text{appearance}}$	1.64	86.38
3	C_1	1.34	86.68
4	$C_{\text{deviation}}$	1.32	86.70
5	Number of words	1.2	86.82
6	C_{stance}	1.13	86.89
7	Lempel-Ziv Complexity	1.08	86.94
8	Alphabet size	1.08	86.94
9	C_2	0.91	87.11
10	C_{arms}	0.55	87.47
11	C_{swing}	0.07	87.95
Reference accuracy		88.02%	
AD/DLB			
Importance Rank	Feature removed	ΔA (in %)	A (in %)
1	C_{arms}	2.54	72.61
2	$C_{\text{step period}}$	1.76	73.39
3	Number of words	0.66	74.49
4	Lempel-Ziv Complexity	0.43	74.72
5	$C_{\text{appearance}}$	0.35	74.80
6	C_{stance}	0.25	74.90
7	C_{swing}	0.19	74.96
8	C_1	-0.04	75.19
9	Alphabet size	-0.51	75.66
10	C_2	-0.90	76.05
11	$C_{\text{deviation}}$	-1.21	76.36
Reference accuracy		75.15%	

Table V: **Feature ablation test results.** A: median classification accuracy. The accuracy difference of feature i is computed to $\Delta A_i = A_{\text{reference}} - A_{\text{without feature } i}$ and reflects the importance of that feature in the classification.