Budapest University of Technology and Economics

Department of Cognitive Science

PhD School in Psychology

Regina Julia Meszlényi

New approaches for fMRI functional connectivity analysis based on Dynamic Time Warping and machine learning

PhD Thesis

Thesis Booklet

Supervisor: Prof. Zoltán Vidnyánszky

Budapest, 2017

Synopsis

Conventional resting-state network concept is based on calculating linear dependence of spontaneous low frequency fluctuations of the BOLD signals of different brain areas (Biswal et al. 1995,Biswal et al. 2010), which assumes temporally stable zero-lag synchrony across regions (Biswal et al. 1995; Fox et al. 2005; Fox and Raichle 2007; Kalcher et al. 2012; Margulies et al. 2010; Yeo et al. 2011). However, growing amount of experimental findings suggest that functional connectivity exhibits dynamic changes and a complex time-lag structure, which cannot be captured by the static zero-lag correlation analysis (Allen et al., 2014; Chang and Glover, 2010; Handwerker et al., 2012; Jones et al., 2012; Kiviniemi et al., 2011; Sakoğlu et al., 2010; Smith, 2012).

In our first study we proposed a new approach applying Dynamic Time Warping distance (Sakoe and Chiba, 1978) to evaluate functional connectivity strength that accounts for non-stationarity and phase-lags between the observed signals, which arises from the dynamic switching between brain states (Allen et al., 2014; Chen et al., 2015). In contrast to well-known methods analysing dynamic functional connectivity (Allen et al., 2014; Hutchison et al., 2013), the DTW algorithm provides a single scalar measure of connectivity strength between brain regions with complex unstable non-zero temporal lag structure, therefore DTW distance and the derived DTW similarity can characterize a wide range of connections while still allowing simple multi-subject statistics. In the first study we examined the applicability of DTW distance for fMRI data analysis and investigated its robustness between measurements and for different preprocessing pipelines.

Resting-state fMRI connectivity based classification gained substantial popularity in the past decade, and raised the intriguing possibility of application of machine learning for fast and objective diagnosis of mental disorders (Abraham et al., 2017; Kassraian-Fard et al., 2016; Arbabshirani et al., 2013; Kim et al., 2016; Rosa et al., 2015; Liem et al., 2017). Connectome based classification studies tend to suffer from the so called 'curse of dimensionality' (Hughes, 1968), i.e. the number of resting-state fMRI measurements in most studies is much lower than the number of pairwise connectivity descriptors between brain regions. Therefore the vast majority of these classification experiments use machine learning techniques that can cope with low sample numbers. One particularly interesting model is the Least Absolute Shrinkage and Selection Operator (LASSO) that performs feature selection during classification, i.e. it selects brain region pairs with altered connectivity between subject groups (Ng et al., 2012; Rosa et al., 2015; Ryali et al., 2010; Tibshirani, 1996). In our second study we examined the DTW distance's sensitivity for differences between subject groups using a connectivity based classification task, and we analysed set of connectivity features selected by means of a LASSO model.

As recent advances in resting-state fMRI based dynamic functional connectivity analysis shows, the time-course of phase relationship of distinct brain regions can reveal information about the switching between brain states (Allen et al., 2014; Chang and Glover, 2010; Chen et al., 2015) and the stability of these phase-relationships can differentiate between subject groups (Córdova-Palomera et al., 2017; Demirtaş et al., 2016; Glerean et al., 2012). Based on these results in our third study we investigated whether the length of the warping path – a simple measure derived from the DTW algorithm that can be used as a proxy for connection stability – can differentiate between subject groups in a LASSO classification paradigm.

In the last couple of years the focus of machine learning research has shifted to deep learning techniques (Krizhevsky et al., 2012). Recent developments in the theory of deep learning demonstrates that in case of high-dimensional datasets with complex structure, deep neural networks have an exponential gain in efficiency over conventional machine learning models, i.e. from the same amount of training data, deep neural networks can learn exponentially more com plex output function (Bengio et al., 2005, 2013; LeCun et al., 2015; Montúfar et al., 2014). Therefore deep models show great potential in fMRI based classification (Kawahara et al., 2017; Kim et al., 2016; Plis et al., 2014; Vieira et al., 2017). In our fourth study we investigated the potential of a deep convolutional neural network architecture to classify subject groups based on combined information from different connectivity descriptors, namely DTW distance and warping path length.

The dissertation addresses the following research questions:

1. *Can we use Dynamic Time Warping distance of resting-state fMRI signals to describe functional connectivity strength?*
2. *Can Dynamic Time Warping distance as a descriptor of functional connectivity discriminate subject groups?*
3. *Can warping path length, a descriptor of functional connection stability discriminate subject groups?*
4. *Dynamic Time Warping distance and warping path length describe fundamentally different properties of functional connectivity. How can we efficiently combine information of the two metrics to increase accuracy in connectome classification tasks?*

Thesis points

Thesis point I: Dynamic Time Warping distance and the derived similarity measure can be efficiently used for resting-state fMRI based functional connectivity strength estimation as it can capture dynamic interactions of brain regions and is more robust against global noise linearly combined with the signal in the BOLD time-series than the more conventional approaches based on linear correlation.

In the first study (Meszlényi et al., 2017b), we investigated a new method for functional connectivity strength calculation based on the Dynamic Time Warping algorithm that applies nonlinear warping on the compared time-series to correct for the dynamically changing phase-lag structure of the measured activity of brain regions. Using simulated fMRI data we found that DTW captures dynamic interactions and it is less sensitive to global noise linearly combined with the signal as compared to conventional correlation analysis. We tested our method using resting-state fMRI data from repeated measurements of a single subject and showed that DTW analysis results in more stable connectivity patterns by reducing the within-subject variability and increasing robustness for preprocessing strategies.

Thesis point II: Functional connectivity strength characterized by Dynamic Time Warping distance is sensitive for group differences, i.e. in resting-state fMRI classification tasks, classifiers trained on connectivity features based on Dynamic Time Warping distance systematically outperformed models based on conventional correlation coefficient features.

In the second study (Meszlényi et al., 2016b), we examined the sensitivity of the functional connectivity strength estimates based on Dynamic Time Warping distance in a classification experiment on a dataset of young adults with and without and attention deficit hyperactivity disorder (ADHD). Gender and ADHD diagnosis were used as classification targets. We applied a LASSO classifier that includes feature selection, i.e. based on the LASSO model we can determine which brain connections and regions differentiate well. Our results showed that functional connectivity networks resulting from DTW-based estimates as compared to the correlation-based ones are more stable and achieve higher classification performance (in terms of averaged F-measure) in both gender and ADHD classification.

Thesis point III: Warping path length, a measure of connection stability derived from the Dynamic Time Warping algorithm contains valuable information about the dynamic properties of the functional connection between brain regions beside functional connectivity strength, therefore warping path length can efficiently be used as connectivity feature in resting-state fMRI based classification tasks.

In the third study (Meszlényi et al., 2016a) we demonstrated further advantages of the DTW algorithm: beside the DTW distance, the algorithm generates the warping path, i.e. the time-delay alignment between the two time-series. On the dataset of young adults we used the length of the warping path as classification feature for cannabis addiction classification and demonstrated that the warping path itself carries important information, as classifiers based on relative warping-path length significantly outperformed models based on either correlation or DTW distance. As we applied LASSO classifiers, we could also demonstrate the network of brain regions that differentiates well between addicted and non-addicted subjects based on the stability of connection between regions.

Thesis point IV: Classification based on Dynamic Time Warping distance and path length combined with a convolutional neural network specifically designed for connectome based classification achieves substantially better performance than classification based on single connectivity metric features (i.e. Dynamic Time Warping distance, warping path length or correlation coefficients alone).

In the fourth study (Meszlényi et al., 2017a) we designed a convolutional neural network architecture specifically for classification based on functional connectivity and tested the method on a datasets of patients suffering from mild cognitive impairment. A great advantage of the convolutional architecture is that it can straightforwardly combine heterogeneous connectivity metrics and can successfully utilize this heterogeneity in contrast to conventional machine learning models. With this convolutional model we demonstrated that the best classification performance can be achieved by combining connectivity information extracted from DTW distance and warping path length features. We also showed that based on a relatively simple analysis of the learnt weights of the convolutional neural network we can determine which brain regions have most influence on the classification output, i.e. the regions that show altered connectivity between groups.

Discussion

To determine the usefulness of a functional connectivity metric we have to take certain properties into account. Connectivity strength between brain regions should be reliable between multiple measurements and also robust against common noise sources, while it should retain sensitivity for differences between subject groups caused for example by mental disorders or other phenotypic differences. Conventional correlation based functional connectivity calculation holds many of these properties, however we hypothesized that a method that can account for the dynamic nature of connections could be even more advantageous.

We were able to prove that as the Dynamic Time Warping algorithm can handle non-stationary processes, the derived connectivity strength results in more stable connectivity patterns in repeated measurements, and is less sensitive for linearly combined common noise than connectomes calculated with correlation coefficients. We could also show that DTW distance differentiates better between groups than correlation as concluded from results of multiple classification studies performed using several different classification algorithms and with different classification targets.

Besides functional connectivity strength calculation, the DTW algorithm also provides information about the time-varying phase difference between brain regions through the generation of the warping path. The length of this warping path can be used to describe the stability of connections between brain regions and it is also sensitive for group differences (in terms of averaged F-measure) as our classification studies show. We were able to demonstrate that the warping path length holds complementary information compared to connectivity strength based on DTW distance, therefore a carefully designed convolutional neural network architecture that is able to straightforwardly combine different connectome descriptors can discriminate subject groups substantially better if it is provided with both DTW distance and warping path length features, than a network trained on single connectivity descriptors.

With these results we were able to demonstrate that the DTW algorithm is indeed an applicable and advantageous tool of resting-state functional connectivity analysis, as it can take into account the dynamic properties of the connections and simultaneously provide functional connectivity strength and connection stability information.

Publications attached to the thesis points:

1. Meszlényi, R. J., Hermann, P., Buza, K., Gál, V., and Vidnyánszky, Z. (2017b). Resting state fMRI functional connectivity analysis using Dynamic Time Warping. *Front. Neurosci.* 11. doi:10.3389/fnins.2017.00075.

2. Meszlényi, R., Peska, L., Gál, V., Vidnyánszky, Z., and Buza, K. (2016b). Classification of fMRI data using dynamic time warping based functional connectivity analysis. in *2016 24th European Signal Processing Conference (EUSIPCO)* (Budapest), 245–249. doi:10.1109/EUSIPCO.2016.7760247.

3. Meszlényi, R., Peska, L., Gál, V., Vidnyánszky, Z., and Buza, K. (2016a). A model for classification based on the functional connectivity pattern dynamics of the brain. in *2016 Third European Network Intelligence Conference (ENIC)* (Wrocław), 203–208. doi:10.1109/ENIC.2016.037.

4. Meszlényi, R., Buza, K., and Vidnyánszky, Z. (2017a). Resting State fMRI Functional Connectivity-Based Classification Using a Convolutional Neural Network Architecture. *Front. Neuroinformatics* 11. doi:10.3389/fninf.2017.00061

References

Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., et al. (2017). Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. *NeuroImage* 147, 736–745. doi:10.1016/j.neuroimage.2016.10.045.

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex N. Y. N 1991* 24, 663–676. doi:10.1093/cercor/bhs352.

Arbabshirani, M. R., Kiehl, K., Pearlson, G., and Calhoun, V. D. (2013). Classification of schizophrenia patients based on resting-state functional network connectivity. *Front. Neurosci.* 7. doi:10.3389/fnins.2013.00133.

Bengio, Y., Courville, A., and Vincent, P. (2013). Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 1798–1828. doi:10.1109/TPAMI.2013.50.

Bengio, Y., Delalleau, O., and Roux, N. L. (2005). The Curse of Highly Variable Functions for Local Kernel Machines. in *Proceedings of the 18th International Conference on Neural Information Processing Systems* NIPS'05. (Cambridge, MA, USA: MIT Press), 107–114. Available at: http://dl.acm.org/citation.cfm?id=2976248.2976262 [Accessed June 2, 2017].

Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., et al. (2010). Toward discovery science of human brain function. *Proc. Natl. Acad. Sci.* 107, 4734–4739. doi:10.1073/pnas.0911855107.

Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541.

Chang, C., and Glover, G. H. (2010). Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage* 50, 81–98. doi:10.1016/j.neuroimage.2009.12.011.

Chen, J. E., Chang, C., Greicius, M. D., and Glover, G. H. (2015). Introducing co-activation pattern metrics to quantify spontaneous brain network dynamics. *NeuroImage* 111, 476–488. doi:10.1016/j.neuroimage.2015.01.057.

Córdova-Palomera, A., Kaufmann, T., Persson, K., Alnæs, D., Doan, N. T., Moberget, T., et al. (2017). Disrupted global metastability and static and dynamic brain connectivity across individuals in the Alzheimer's disease continuum. *Sci. Rep.* 7, 40268. doi:10.1038/srep40268.

Demirtaş, M., Tornador, C., Falcón, C., López-Solà, M., Hernández-Ribas, R., Pujol, J., et al. (2016). Dynamic functional connectivity reveals altered variability in functional connectivity among patients with major depressive disorder. *Hum. Brain Mapp.* 37, 2918–2930. doi:10.1002/hbm.23215.

Fox, M. D., and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711. doi:10.1038/nrn2201.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Essen, D. C. V., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9673–9678. doi:10.1073/pnas.0504136102.

Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., and Sams, M. (2012). Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connect.* 2, 91–101. doi:10.1089/brain.2011.0068.

Handwerker, D. A., Roopchansingh, V., Gonzalez-Castillo, J., and Bandettini, P. A. (2012). Periodic changes in fMRI connectivity. *NeuroImage* 63, 1712–1719. doi:10.1016/j.neuroimage.2012.06.078.

Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Inf. Theory* 14, 55–63. doi:10.1109/TIT.1968.1054102.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage* 80, 360–378. doi:10.1016/j.neuroimage.2013.05.079.

Jones, D. T., Vemuri, P., Murphy, M. C., Gunter, J. L., Senjem, M. L., Machulda, M. M., et al. (2012). Non-Stationarity in the "Resting Brain's" Modular Architecture. *PLoS ONE* 7. doi:10.1371/journal.pone.0039731.

Kalcher, K., Huf, W., Boubela, R. N., Filzmoser, P., Pezawas, L., Biswal, B., et al. (2012). Fully exploratory network independent component analysis of the 1000 functional connectomes database. *Front. Hum. Neurosci.* 6, 301. doi:10.3389/fnhum.2012.00301.

Kassraian-Fard, P., Matthis, C., Balsters, J. H., Maathuis, M. H., and Wenderoth, N. (2016). Promises, Pitfalls, and Basic Guidelines for Applying Machine Learning Classifiers to Psychiatric Imaging Data, with Autism as an Example. *Front. Psychiatry* 7. doi:10.3389/fpsyt.2016.00177.

Kawahara, J., Brown, C. J., Miller, S. P., Booth, B. G., Chau, V., Grunau, R. E., et al. (2017). BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage* 146, 1038–1049. doi:10.1016/j.neuroimage.2016.09.046.

Kim, J., Calhoun, V. D., Shim, E., and Lee, J.-H. (2016). Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *NeuroImage* 124, Part A, 127–146. doi:10.1016/j.neuroimage.2015.05.018.

Kiviniemi, V., Vire, T., Remes, J., Elseoud, A. A., Starck, T., Tervonen, O., et al. (2011). A sliding time-window ICA reveals spatial variability of the default mode network in time. *Brain Connect.* 1, 339–347. doi:10.1089/brain.2011.0036.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, eds. F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc.), 1097–1105. Available at: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf [Accessed March 7, 2017].

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi:10.1038/nature14539.

Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Kharabian Masouleh, S., Huntenburg, J. M., et al. (2017). Predicting brain-age from multimodal imaging data captures cognitive impairment. *NeuroImage* 148, 179–188. doi:10.1016/j.neuroimage.2016.11.005.

Margulies, D. S., Böttger, J., Long, X., Lv, Y., Kelly, C., Schäfer, A., et al. (2010). Resting developments: a review of fMRI post-processing methodologies for spontaneous brain activity. *Magn. Reson. Mater. Phys. Biol. Med.* 23, 289–307. doi:10.1007/s10334-010-0228-5.

Meszlényi, R. J., Buza, K., and Vidnyánszky, Z. (2017a). Resting State fMRI Functional Connectivity-Based Classification Using a Convolutional Neural Network Architecture. *Front. Neuroinformatics* 11. doi:10.3389/fninf.2017.00061.

Meszlényi, R. J., Hermann, P., Buza, K., Gál, V., and Vidnyánszky, Z. (2017b). Resting State fMRI Functional Connectivity Analysis Using Dynamic Time Warping. *Front. Neurosci.* 11. doi:10.3389/fnins.2017.00075.

Meszlényi, R., Peska, L., Gál, V., Vidnyánszky, Z., and Buza, K. (2016a). A Model for Classification Based on the Functional Connectivity Pattern Dynamics of the Brain. in *2016 Third European Network Intelligence Conference (ENIC)* (Wrocław), 203–208. doi:10.1109/ENIC.2016.037.

Meszlényi, R., Peska, L., Gál, V., Vidnyánszky, Z., and Buza, K. (2016b). Classification of fMRI data using dynamic time warping based functional connectivity analysis. in *2016 24th European Signal Processing Conference (EUSIPCO)* (Budapest), 245–249. doi:10.1109/EUSIPCO.2016.7760247.

Montúfar, G., Pascanu, R., Cho, K., and Bengio, Y. (2014). On the Number of Linear Regions of Deep Neural Networks. in *Proceedings of the 27th International Conference on Neural Information Processing Systems* NIPS'14. (Cambridge, MA, USA: MIT Press), 2924–2932. Available at: http://dl.acm.org/citation.cfm?id=2969033.2969153 [Accessed June 2, 2017].

Ng, B., Siless, V., Varoquaux, G., Poline, J. B., Thirion, B., and Abugharbieh, R. (2012). Connectivity-informed Sparse Classifiers for fMRI Brain Decoding. in *2012 International Workshop on Pattern Recognition in NeuroImaging (PRNI)*, 101–104. doi:10.1109/PRNI.2012.11.

Plis, S. M., Hjelm, D. R., Salakhutdinov, R., Allen, E. A., Bockholt, H. J., Long, J. D., et al. (2014). Deep learning for neuroimaging: a validation study. *Front. Neurosci.* 8. doi:10.3389/fnins.2014.00229.

Rosa, M. J., Portugal, L., Hahn, T., Fallgatter, A. J., Garrido, M. I., Shawe-Taylor, J., et al. (2015). Sparse network-based models for patient classification using fMRI. *NeuroImage* 105, 493–506. doi:10.1016/j.neuroimage.2014.11.021.

Ryali, S., Supekar, K., Abrams, D. A., and Menon, V. (2010). Sparse logistic regression for whole-brain classification of fMRI data. *NeuroImage* 51, 752–764. doi:10.1016/j.neuroimage.2010.02.040.

Sakoe, H., and Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. Acoust. Speech Signal Process.* 26, 43–49. doi:10.1109/TASSP.1978.1163055.

Sakoğlu, Ü., Pearlson, G. D., Kiehl, K. A., Wang, Y. M., Michael, A. M., and Calhoun, V. D. (2010). A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia. *Magma N. Y. N* 23, 351–366. doi:10.1007/s10334-010-0197-8.

Smith, S. M. (2012). The future of FMRI connectivity. *NeuroImage* 62, 1257–1266. doi:10.1016/j.neuroimage.2012.01.022.

Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* 106, 1125–1165. doi:10.1152/jn.00338.2011.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* 58, 267–288.

Vieira, S., Pinaya, W. H. L., and Mechelli, A. (2017). Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neurosci. Biobehav. Rev.* 74, Part A, 58–75. doi:10.1016/j.neubiorev.2017.01.002.