Kernel-based Nonlinear Manifold Learning for EEG Functional Connectivity Analysis with Application to Alzheimer's Disease

Gunawardena, S. R. A. S., Sarrigiannis, P. G., Blackburn, D. J. & He, F

Original citation & hyperlink:
Gunawardena, SRAS, Sarrigiannis, PG, Blackburn, DJ & He, F 2023, Kernel-based Nonlinear Manifold Learning for EEG Functional Connectivity Analysis with Application to Alzheimer's Disease. in 2022 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). IEEE, pp. 1-5.
https://doi.org/10.1109/spmb55497.2022.10014969

DOI 10.1109/spmb55497.2022.10014969
ISBN 978-1-6654-7030-8
ESBN 978-1-6654-7030-8

Publisher: IEEE

Kernel-based Nonlinear Manifold Learning for EEG Functional Connectivity Analysis with Application to Alzheimer's Disease

R. Gunawardena[1], P. G. Sarrigiannis[2], D. J. Blackburn[3] and F. He[1]

1. Centre for Computational Science and Mathematical Modelling, Coventry University, Coventry, UK
2. Department of Neurophysiology, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK
3. Department of Neuroscience, University of Sheffield, Sheffield, UK
{gunawardes, fei.he}@coventry.ac.uk, ptolemaios.sarrigiannis@nhs.net, d.blackburn@sheffield.ac.uk

I. Introduction

Dynamical , causal and cross-frequency coupling analysis using the EEG has received significant interest for the analysis and diagnosis of neurological disorders [1, 2, 3]. Due to the high computational requirements needed for some of these methods, EEG channel selection is crucial [4]. Functional connectivity (FC) between EEG channels is often used for channel selection and connectivity analysis [4, 5, 6]. Ideally, in the case of selecting channels for dynamical and causal analysis, FC methods should be able to account for linear and nonlinear spatial and temporal interactions between EEG channels. In neuroscience, FC is quantified using different measures of (dis)similarity to assess the statistical dependence between two signals [5]. However, the interpretation of FC measures can differ significantly from one measure to another[5, 7]. In the early diagnosis of AD, [7] showed correlations among various (dis)similarity measures, and therefore these measures can be grouped. Thus, one from each is sufficient to extract information from the data [7]. Therefore, the development of a generic measure of (dis)similarity is important in FC analysis.

To learn the spatial and temporal structures within the data, in this study, kernel-based nonlinear manifold learning is used. We introduce a novel EEG FC analysis method to determine a subset of important channel inter-relationships for dynamical and causal analysis. For instance the changes due to neurodegeneration, e.g. Alzheimer's disease (AD), on global and local brain dynamics. Our FC analysis method uses robust kernel Isomap as initialisation to the latent space for the Gaussian Process Latent Variable Model (Isomap-GPLVM). The Isomap-GPLVM method is used to learn the spatio-temporal local and global (dis)similarities present within the EEG data. The resulting kernel matrix quantifies this information as a generic measure of (dis)similarity and can be used to assess linear and nonlinear FC, between EEG channels.

In this work, EEG data from healthy controls (HC) and patients with mild to moderate AD are used as a case study. Kernel (dis)similarity matrices for each participant are evaluated and linear SVM classification is used to assess how well our FC measure can differentiate between HC and AD groups. FC analysis for both eyes-open (EO) and eyes-closed (EC) conditions are shown. Classification performance between our measure and other commonly used FC measures is highlighted to show the efficacy of our method.

II. Data

The specifics of the experimental design, diagnosis, data acquisition and EEG electrode placements are provided in [8]. In summary; three 12-second epochs of EO/EC EEGs from 20 HC and 20 AD participants are used in this study. The following 23 bipolar channels, from a 10-20 international electrode arrangement, were used for this work: F8–F4, F7–F3, F4–C4, F3–C3, F4–FZ, FZ–CZ, F3–FZ, T4–C4, T3–C3, C4–CZ, C3–CZ, CZ–PZ, C4–P4, C3–P3, T4–T6, T3–T5, P4–PZ, P3–PZ, T6–O2, T5–O1, P4–O2, P3–O1 and O2–O1. EEG frequencies between 2 to 100Hz are only used. Frequency components around 50Hz (49.5-50.5Hz) were removed to avoid any contamination by AC power line noise. Unwanted frequency components were removed using the fast Fourier transform (FFT) and inverse-FFT. Time-domain signals were then down-sampled from 2kHz to 200Hz. The EEG data are then normalised to zero mean and unit variance.

978-1-6654-7029-2/22/$31.00 ©2022 IEEE IEEE SPMB 2022 December 3, 2022

III. METHODS

Manifold learning methods learn local similarities or global dissimilarities within the data and embed this information in the lower-dimensional space (or latent space) [9]. Methods that imply a data-to-latent space mapping, such as Isomap and KernelPCA, preserve local similarities in the latent space[9]. Conversely, methods such as GPLVM, which imply a latent-to-data space mapping, preserve global dissimilarities [9]. In kernel-based manifold learning methods, the kernel describes the learnt linear/nonlinear structures within the data in the form of pairwise (dis)similarity. Therefore, such methods can be used to obtain a lower-dimensional representation of spatio-temporal structures within EEG data. Through temporal dimensional reduction, the kernel matrix obtained will reflect this information as a measure of (dis)similarity between EEG channels and will be named a kernel (dis)similarity matrix.

The methodology presented in this work compromises between learning local similarities and global dissimilarities. This is achieved by using robust kernel Isomap [10] to initialise GPLVM; we name it Isomap-GPLVM. Robust kernel Isomap approximates the geodesic distance to obtain a lower-dimensional representation of the data, preserving local similarities and providing a method to eliminate critical outliers [10, 9]. The choice of the latent dimension determines the size of the lower-dimensional representation after dimensionality reduction. GPLVM, a kernel-based probabilistic latent variable model, uses the Isomap representation of the data as an initialisation of the latent space. GPLVM then operates on this latent space according to the distance between the high-dimensional data points, preserving global dissimilarities [9]. GPLVM achieves this by maximising the marginal likelihood of the data space with respect to the latent space and the kernel hyper-parameters. The latent space to data space mapping in GPLVM is a function of the kernel matrix [11] and determines the regions of similarity and dissimilarity between latent variables [12]. The kernel matrix contains both local similarity and global dissimilarity information that is learnt from the data and will be called the kernel (dis)similarity matrix. The Radial Basis Function (RBF) is used as the kernel function for GPLVM. Therefore, the kernel matrix will represent the local/global (dis)similarities learnt as a generic measure of similarity. The RBF kernel is used as the kernel function for GPLVM as it has the universal approximation property and can be integrated into most functions to achieve smooth mapping from latent space to data space [12].

Reducing the temporal dimension in each EEG data, participant-specific kernel (dis)similarity matrices of the HC and AD groups are evaluated using Isomap-GPLVM for all three epochs. The choice of the latent dimension and the initial conditions for the hyper-parameters for the RBF kernel are chosen (via grid search) depending on how well the pairwise features from HC and AD kernel (dis)similarity matrices can distinguish between the groups. Linear SVM with Monte-Carlo cross-validation (SVM-MCV) is used to assess this. The initial conditions that result in the highest area under the receiver operator curve (AU-ROC) are chosen. The training set comprises 10 HC and 10 AD kernel matrices randomly picked from the first epoch. The testing set is composed of the remaining 10 HC and 10 AD kernel matrices from the first epoch. The testing set will also include all the HC/AD kernel matrices from the second and third epochs. 1000 such random samples are taken to generate several testing and training sets. The average AU-ROC from the testing sets is used to assess the performance. Due to the small data set, this cross-validation strategy is deemed appropriate to obtain a fair balance.

IV. RESULTS

The 23 bipolar montage EEG channels used can be grouped according to the respective underlying cortical regions as follows; Occipital: O1-O2, P4-O2, P3-O1, T5-O1, T6-O2. Parietal: P3-PZ, P4-PZ. Temporal: T3-T5, T4-T6. Centro-Parietal: C3-P3, C4-P4, CZ-PZ, C3-CZ, C4-CZ. Centro-Temporal: T3-C3, T4-C4. Fronto-Central: F3-C3, FZ-CZ, F3-FZ, F4-FZ, F4-C4. Frontal: F7-F3, F8-F4. Fronto-Central and Centro-Parietal regions can be grouped as Fronto-Parietal. EEG is a scalp level sensor with a low spatial resolution. It does not measure the actual activity of the cerebral cortex. Cortical area name is used as a marker.

The AU-ROCs of our kernel (dis)similarity based FC measure are 0.73 ± 0.04 and 0.77 ± 0.05 (EO and EC respectively). Applying SVM-MCV to commonly used FC measures, the AU-ROCs for correlation are 0.67 ± 0.04 and 0.61 ± 0.03. For phase-lag value (PLV) 0.70 ± 0.04 and 0.76 ± 0.04. In the above comparisons the the former AU-ROC stated is for the EO case and the next is for the EC case.

The Mann–Whitney U test is applied for statistical comparison and the Benjamini-Hochberg false discovery rate controlling method to mitigate the problem of multiple statistical comparisons. This was applied element-wise to the resulting kernel (dis)similarity matrices between the groups. This information is illustrated in Figure 1 in the form of a matrix where blue indicates the statistically significant EEG channel pairwise (dis)similarities. Both conditions EO and EC are shown. A summary of the comparisons is given below.

Common statistically significant FC changes in both conditions–between channels from the regions; Parietal and Occipital, Parietal and Fronto-Parietal, Fronto-Central and Centro-Parietal. However, these pairwise FC changes had low average weightings in SMV-MCV. A notable observation that had a higher weighting in both cases is the significant FC changes between channel T6-02 (index 4, located in the bottom right hemisphere) and the channels from regions Parietal, Fronto-Parietal, and Frontal.

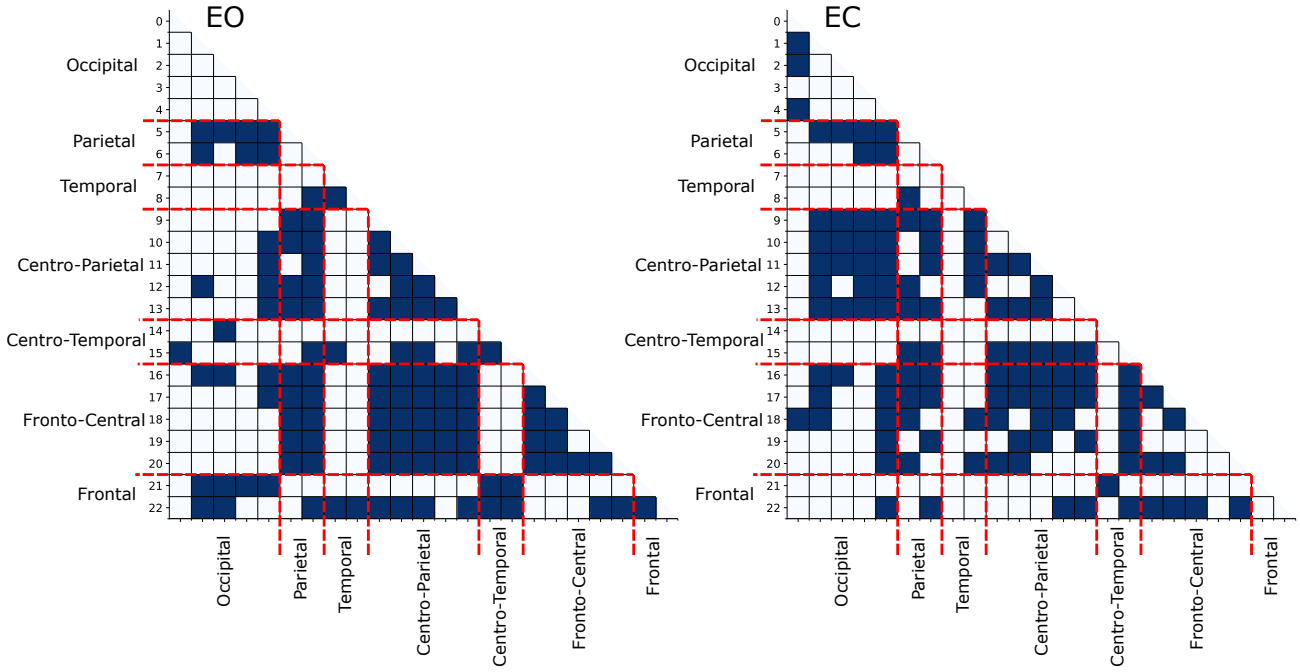

Figure 1. Statistically significant group differences in EEG FC – comparison of EO and EC. Blue indicates significant change in the respective pairwise channel FC between groups

V. Conclusions and Future Work

This study presents a novel EEG channel FC measure. Kernel-based manifold learning was used to learn (dis)similarity information within the EEG data to determine important FC changes between pairs of EEG channels, in the case of AD. Compromising between global and local spatio-temporal structures within the EEG data for linear and nonlinear FC analysis is a main feature of our method. Following this work, a more comprehensive FC analysis and comparisons with a range of common FC measures are required. By including more participant data, the methodology can potentially be used as a diagnostic tool for neurodegenerative diseases, which not only gives a classification but also provides an analysis of FC changes.

ACKNOWLEDGEMENTS

RG and FH acknowledge Coventry University for the Trailblazer PhD studentship. The EEG data was funded by a grant from Alzheimer's Research UK, grant reference number ARUK-PPG20114B-25. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

[1] C. Babiloni, R. Lizio, N. Marzano, P. Capotosto, A. Soricelli, A. I. Triggiani, S. Cordone, L. Gesualdo, and C. Del Percio, "Brain neural synchronization and functional coupling in alzheimer's disease as revealed by resting state EEG rhythms," *International Journal of Psychophysiology*, vol. 103, pp. 88–102, 2016, *r*esearch on Brain Oscillations and Connectivity in A New Take-Off State.

[2] O. Jensen, E. Spaak, and J. M. Zumer, "Human brain oscillations: From physiological mechanisms to analysis and cognition," in *Magnetoencephalography: From Signals to Dynamic Cortical Networks*, S. Supek and C. J. Aine, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 359–403.

[3] F. He and Y. Yang, "Nonlinear system identification of neural systems from neurophysiological signals," *Neuroscience*, vol. 458, pp. 213–228, 2021.

[4] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, "A review of channel selection algorithms for EEG signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2015, no. 1, p. 66, Aug 2015.

[5] R. Mohanty, W. A. Sethares, V. A. Nair, and V. Prabhakaran, "Rethinking measures of functional connectivity via feature extraction," *Scientific Reports*, vol. 10, no. 1, p. 1298, Jan 2020.

[6] S. Sadaghiani, M. J. Brookes, and S. Baillet, "Connectomics of human electrophysiology," *NeuroImage*, vol. 247, p. 118788, 2022.

[7] J. Dauwels, F. Vialatte, T. Musha, and A. Cichocki, "A comparative study of synchrony measures for the early diagnosis of alzheimer's disease based on EEG," *NeuroImage*, vol. 49, no. 1, pp. 668–693, 2010.

[8] D. Blackburn, Y. Zhao, M. D. Marco, S. Bell, F. He, H.-L. Wei, S. Lawrence, Z. Unwin, M. Blyth, J. Angel, K. Baster, T. Farrow, I. Wilkinson, S. Billings, A. Venneri, and P. Sarrigiannis, "A pilot study investigating a novel non-linear measure of eyes open versus eyes closed EEG synchronization in people with alzheimer's disease and healthy controls," *Brain Sciences*, vol. 8, no. 7, p. 134, jul 2018.

[9] N. D. Lawrence and J. Quiñonero Candela, "Local distance preservation in the gp-lvm through back constraints," in *Proceedings of the 23rd International Conference on Machine Learning*, ser. ICML '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 513–520.

[10] H. Choi and S. Choi, "Robust kernel isomap," *Pattern Recognition*, vol. 40, no. 3, pp. 853–862, 2007.

[11] N. D. Lawrence, "Gaussian process latent variable models for visualisation of high dimensional data," in *Proceedings of the 16th International Conference on Neural Information Processing Systems*, ser. NIPS'03. Cambridge, MA, USA: MIT Press, 2003, p. 329–336.

[12] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2005.