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

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

Original citation:
Gunawardena, R, Sarrigiannis, PG, Blackburn, DJ & He, F 2023, 'Kernel-based Nonlinear Manifold Learning for EEG-based Functional Connectivity Analysis and Channel Selection with Application to Alzheimer's Disease', Neuroscience, vol. 523, pp. 140-156.
https://dx.doi.org/10.1016/j.neuroscience.2023.05.033

DOI 10.1016/j.neuroscience.2023.05.033
ISSN 0306-4522
ESSN 1873-7544

Publisher: Elsevier

NEUROSCIENCE
RESEARCH ARTICLE
R. Gunawardena et al. / Neuroscience 523 (2023) 140–156

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

Rajintha Gunawardena, [a] **Ptolemaios G. Sarrigiannis,** [b] **Daniel J. Blackburn** [c] **and Fei He** [a]*

[a] *Centre for Computational Science and Mathematical Modelling, Coventry University, Coventry CV1 5FB, UK*

[b] *Department of Neurophysiology, Royal Devon and Exeter, NHS Foundation Trust, Exeter EX2 5DW, UK*

[c] *Department of Neuroscience, The University of Sheffield, Sheffield S10 2HQ, UK*

Abstract—Dynamical, causal, and cross-frequency coupling analysis using the electroencephalogram (EEG) has gained significant attention for diagnosing and characterizing neurological disorders. Selecting important EEG channels is crucial for reducing computational complexity in implementing these methods and improving classification accuracy. In neuroscience, measures of (dis) similarity between EEG channels are often used as functional connectivity (FC) features, and important channels are selected via feature selection. Developing a generic measure of (dis) similarity is important for FC analysis and channel selection. In this study, learning of (dis) similarity information within the EEG is achieved using kernel-based nonlinear manifold learning. The focus is on FC changes and, thereby, EEG channel selection. Isomap and Gaussian Process Latent Variable Model (Isomap-GPLVM) are employed for this purpose. The resulting kernel (dis) similarity matrix is used as a novel measure of linear and nonlinear FC between EEG channels. The analysis of EEG from healthy controls (HC) and patients with mild to moderate Alzheimer's disease (AD) are presented as a case study. Classification results are compared with other commonly used FC measures. Our analysis shows significant differences in FC between bipolar channels of the occipital region and other regions (i.e. parietal, centro-parietal, and fronto-central) between AD and HC groups. Furthermore, our results indicate that FC changes between channels along the fronto-parietal region and the rest of the EEG are important in diagnosing AD. Our results and its relation to functional networks are consistent with those obtained from previous studies using fMRI, resting-state fMRI and EEG.

Key words: Alzheimer's disease, EEG, channel selection, manifold learning, machine learning, functional connectivity.2000 MSC: 0000; 1111.

INTRODUCTION

Electroencephalogram (EEG), recorded at the scalp level, reflects the grossly summed currents of the electrical fields generated by neural activity, in the cortical neural circuits. Thus, through EEG, the behaviour and integrity of the underlying neural circuits can be indirectly studied (Nunez and Srinivasan, 2006; Rodriguez-Bermudez and Garcia-Laencina, 2015). Therefore, analysing hidden structures within EEG data is important and has gained considerable attention (Stam, 2005; Rodriguez-Bermudez and Garcia-Laencina, 2015).

In-depth dynamical analysis, such as the analysis of linear and nonlinear dynamic relationships between EEG channels, causality, and cross-frequency coupling analysis, has received much interest (Rodriguez-Bermudez and Garcia-Laencina, 2015; Stam, 2005; Jensen et al., 2014; He et al., 2014a; He and Yang, 2021; He et al., 2016). However, some of these methods can often incur high computational complexity. Consequently, in practice, to reduce the computational complexity, improve classification accuracy and gain prior knowledge on which underlying cortical regions might be important in AD, the selection of important EEG channels from high dimensional EEG data is vital (Alotaiby et al., 2015). Furthermore, to select channels to perform nonlinear dynamical analysis, the channel selection method should be able to account for nonlinear dependencies

*Corresponding author.
E-mail address: fei.he@coventry.ac.uk (F. He).
Abbreviations: AD, Alzheimer's Disease; HC, Healthy Controls; AUROC, Area Under Reciever Operator Curve; EC, Eyes-Closed; EO, Eyes-Open; FC, Functional Connectivity; fMRI, functional Magnetic Resonance Imaging; GPLVM, Gaussian Process latent Variable Model; RBF, Radial Basis Function; SVM, Support Vector Machines; SVM-MCV, SVM Monte-Carlo Cross validation; DAN, Dorsal Attention Network; VAN, Ventral Attention Network; FPN, Fronto Parietal Network; DMN, Default Mode Network.

https://doi.org/10.1016/j.neuroscience.2023.05.033
0306-4522/

between channels (Rodriguez-Bermudez and Garcia-Laencina, 2015; Stam, 2005).

Several recent studies have demonstrated the usefulness of EEG biomarkers in diagnosing and monitoring the progression of AD (Gallego-Jutglà et al., 2015; Li et al., 2019; Babiloni et al., 2020; Jiao et al., 2023). Individuals with MCI and AD typically exhibit decreased alpha/beta power and increased theta/delta power in various brain regions (Horvath et al., 2018). Abnormal changes in FC measures and entropy have also been observed (Dauwels et al., 2011; Gallego-Jutglà et al., 2015; Maturana-Candelas et al., 2019; Li et al., 2019; Babiloni et al., 2020; Klepl et al., 2022). (Jiao et al., 2023) found specific neural biomarkers associated with cognitive function in AD patients, including changes in the power spectrum of low-frequency oscillations in the occipital area and altered signal complexity in the parietal and occipital regions. They also determined that spectral density features and entropy were key EEG biomarkers in differentiating between HC and patients with AD and mild cognitive impairment.

In neuroscience, FC is used to assess the statistical dependence between brain regions (EEG channels) (Babiloni et al., 2016) and is characterised by different measures of (dis) similarity (Mohanty et al., 2020; Falk et al., 2012; Deng et al., 2017; Tylová et al., 2018; Abásolo et al., 2009; Fraga et al., 2013; Al-Qazzaz et al., 2014; Tzimourta et al., 2019; Dauwels et al., 2011; Briels et al., 2020)–distance measures, entropy, and mutual information (Briels et al., 2020; Mohanty et al., 2020; Dauwels et al., 2010b). Some of these measures can be used to analyse nonlinear structures present locally and globally, within the EEG data (Zerzucha and Walczak, 2012; Dauwels et al., 2010b; Mohanty et al., 2020; Babiloni et al., 2016; Briels et al., 2020; Dauwels et al., 2010a) and are often used in many EEG channel selection approaches (Alotaiby et al., 2015). However, regardless of structural connectivity, brain regions functionally connected under one measure do not necessarily imply the same with another measure, as they could even be disconnected (Dauwels et al., 2010b; Mohanty et al., 2020; Briels et al., 2020). Dauwels et al. (Dauwels et al., 2010b) showed that various (dis) similarity measures could be correlated to each other, such as in the application of early diagnosis of AD. Therefore, these correlated measures can often be grouped, and a measure from each group is sufficient to analyse the structures within the data (Dauwels et al., 2010b). Therefore, developing a generic measure of (dis) similarity is important for analysing brain FC and channel selection (Briels et al., 2020).

Similarity or dissimilarity between two variables (EEG channels), in general, express the degree to which the two objects are respectively alike/related or different/distinct (Laub et al., 2006; Shirkhorshidi et al., 2015). Local similarities refer to the relatedness or correlation between nearby data points. This entails that data points closer together in space, time, or any other relevant dimension tend to have similar values or characteristics. Conversely, global dissimilarities refer to a lack of correlation or differences between data points that are far apart from each other. It is important that FC measures account for both local similarities and global dissimilarities within the EEG (Dauwels et al., 2010b; Mohanty et al., 2020; Briels et al., 2020).

For most EEG channel selection techniques, features from the channels are first extracted, and important channels are selected via feature selection. These feature selection methods can be categorised into the following three groups (Alotaiby et al., 2015). a) *Filtering methods*: Independent evaluation criteria, including FC measures, are used for channel selection. Depending on the criteria, these are often only based on single or pairwise EEG channel(s). Filtering methods are good at eliminating irrelevant and redundant features. b) *Wrapper methods*: Subsets of features are generated based on a method of choice. Each subset is evaluated using a classification algorithm to select a subset of channels. These are based on greedy search algorithms aiming to find the best possible combination of features. c) *Embedded Methods*: These techniques simultaneously perform feature selection and classification. For example–LASSO-based feature selection, logistic regression and decision tree are techniques that come under embedded methods. Ranking of features can be easily done using embedded methods.

In this study, learning spatio-temporal linear and nonlinear (dis) similarities within the data is achieved using kernel-based nonlinear manifold learning (dimensionality reduction). The focus is on the differences in EEG FC between healthy and patient groups including the selection of important EEG channels. The motivation behind using a kernel-based method is, the learnt (dis) similarity information is reflected in the kernel, as a generic measure of distance (Schölkopf, 2000) (pairwise comparisons between EEG channels). In this work, the kernel matrix is evaluated using Gaussian Process Latent Variable Model (GPLVM) (Lawrence, 2003). Robust kernel Isomap (Choi and Choi, 2007) is used as an initialisation method, for GPLVM (Isomap-GPLVM). This enables the learning of both local similarities and global dissimilarities within the EEG data and embedding this information in the reduced-dimension manifold (latent space) (Lawrence and Quiñonero Candela, 2006). Furthermore, since dimensionality reduction is used to reduce the temporal dimension, temporal structures within the data are taken into account in the latent space. Considering the above, the kernel matrix evaluated using Isomap-GPLVM can be regarded as a more objective (dis) similarity measure containing information on both linear and nonlinear spatio-temporal EEG inter-relationships. It is a generalisation of different functional connectivity measures (Schölkopf, 2000) and can be a better alternative to using various (dis) similarity measures (Dauwels et al., 2010b; Mohanty et al., 2020; Briels et al., 2020). Based on this novel FC measure, we introduce an EEG channel selection method to determine which channel inter-relationships are more important for in-depth neural dynamical analysis, such as, understanding the effect of neurodegeneration on global and local brain dynamics. This work presents, the analysis of EEG data from a

cohort of age-matched healthy controls (HC) and patients with mild to moderate AD as a case study.

Participant-specific kernel matrices are obtained using Isomap-GPLVM. Linear SVM classification with Monte-Carlo cross-validation (SVM-MCV) is used to assess, how well the proposed FC measure can differentiate between HC and AD groups. FC analysis is presented for both eyes-open (EO) and eyes-closed (EC) conditions. Linear SVM-MCV is also used to rank the selected pairwise features. Therefore, the proposed channel selection method is a hybrid form (Alotaiby et al., 2015) of the aforementioned categories of feature selection methods. Specifically, the proposed approach is an integration of filtering and embedded methods. The channel pairs chosen using our approach can be linked to other EEG studies in the literature considering connectivity analysis in AD. This study aims to introduce and demonstrate the efficacy of our method by comparing it with other commonly used FC measures.

This paper is organised as follows. Specifics about the EEG data, participants and the pre-processing steps are provided in Section 'Data'. Section 'Experimental procedures' discusses the manifold learning methodology via Isomap-GPLVM and the use of related kernel-based (dis) similarity matrix for the classification of EEG data, which are measured from a group of AD patients and an age-matched healthy control cohort. This section also presents the linear SVM and Monte-Carlo cross-validation procedures. Section 'Results' presents the results obtained followed by a discussion of the results in Section 'Discussion'. Limitations of the study and possible improvements to the methodology are discussed along with the concluding remarks in Section 'Discussion'.

DATA

In this work, we include a total of 20 AD cases and 20 age and gender-matched healthy controls (HC) (less than 70 years of age), which are selected based on clinical and radiological diagnostic criteria as described in (Blackburn et al., 2018). Task-free EEG recordings that require minimal cooperation of AD patients are used; typically, this group of patients can have difficulty engaging with or following cognitive tasks. The details of experimental design, diagnosis confirmation, data acquisition and EEG electrode configuration are provided in (Blackburn et al., 2018). All AD participants were in the mild to moderate stage of the disease at the time of EEG recordings.

The Sheffield Teaching Hospital memory clinic team, focusing mainly on young-onset memory disorder, recruited all AD participants. AD participants were diagnosed between 1 month and 2 years before data collection. The diagnosis was made using a series of psychological tests, medical history, neuro-radiological examinations and neurological examinations. High resolution structural magnetic resonance imaging (MRI) was used to eliminate other causes of dementia in all participants. The age and gender-matched HC participants were recruited, whose structural MRI scans and neuropsychological tests were normal. This study was approved by the Yorkshire and The Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). All participants gave their informed written consent.

EEG data

A modified 10–10 overlapping 10–20 international system of electrode placement method was used to acquire EEG recordings. All EEG data were recorded using an XLTEK 128-channel headbox with Ag/AgCL electrodes placed on the scalp at a sampling frequency of 2 kHz. A common referential montage with linked earlobe reference was used. During the 30 min of EEG recording, participants were encouraged not to think about anything specific. All participants had their eyes-open (EO) for 2 min and then eyes-closed (EC) for 2 min, in repeat, during the 30-min recording. The EEG data were reviewed by an experienced neurophysiologist on the XLTEK review station with time-locked video recordings (Optima Medical LTD). Subsequently, from the resting-state EEG recordings, three 12-s artefact-free epochs under EO and EC conditions were isolated.

From the referential montage, the following 23 bipolar channels are produced for the analysis: 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. The bipolar channels are obtained by simply subtracting the two common referenced signals involved. In summary, three 12-s epochs of EO EEGs are collected from 20 HC and 20 AD participants and used in this study.

It should be noted, that a bipolar montage is preferred in several studies (Falk et al., 2012; Trambaiolli et al., 2011) due to evidence of inter-hemispheric disconnection in patients with AD (Jeong, 2004). Furthermore, Nunez et al. in (Nunez and Srinivasan, 2006) explains in detail that the EEG bipolar montages, given the bipolar electrode pairs are sufficiently close (1–3 cm), can be effective in improving the spatial resolution of the EEG due to better estimation of localised electric fields along the scalp surface. Bipolar channels estimate the instantaneous electric field along the scalp surface midway between the pair of electrodes (Nunez and Srinivasan, 2006). To avoid confusion, from now on any bipolar channel(s) will be referred to as *EEG channel(s)*, or just *channel(s)*, unless otherwise specified.

Pre-processing tasks

In this study, since the high-dimensional temporal structures of the multi-channel EEG are examined, the use of filters would pose a major issue due to the phase-related distortions induced (Luck, 2014). Therefore, we first convert the time-series EEG data to the frequency domain using the Fast Fourier Transform (FFT). Thus, unwanted frequency components can be easily removed with minimal phase distortions. Thereafter, inverse-FFT is used to reconstruct the time-domain signals without the unwanted frequency components. The

analysis in this work is performed using EEG frequencies between 2 and 100 Hz. Frequencies below 2 Hz, are not used to avoid low-frequency artefacts due to eye-blinking and slow movements. Furthermore, frequency components around 50 Hz (49.5–50.5 Hz) are also removed to avoid any contamination by AC power line noise. After removing the unwanted frequency components, the reconstructed time-domain signals are then down-sampled to 200 Hz.

METHODOLOGICAL PROCEDURES

This paper introduces a novel measure of FC, a methodology that employs kernel-based manifold learning to identify important channel inter-relationships (channel pairs) within the EEG data for the case of AD. Manifold learning is a nonlinear dimensionality reduction technique that learns a lower-dimensional representation from high-dimensional data (Lawrence and Quiñonero Candela, 2006). EEG data comprises multi-channel time-series data, which is high-dimensional spatio-temporal data. Kernel-based manifold learning can reduce the temporal dimension and learn both linear and nonlinear spatial and temporal structures within the EEG data. The kernel matrix obtained from such manifold learning methods will contain this information as a measure of (dis) similarity and will be named a kernel (dis) similarity matrix. This matrix can be used as a general measure of spatio-temporal functional connectivity.

Manifold learning techniques that maintain local similarities in the lower-dimensional space (also called latent space) entail a mapping from the data space to the latent space (Lawrence, 2005; Lawrence and Quiñonero Candela, 2006). This ensures that data points relatively close in the data space are positioned close together in the latent space. Kernel principal component analysis (KPCA), locally linear embedding (LLE), t-SNE, and Isomap are examples of such techniques. In contrast, techniques that involve a mapping from the latent space to the data space preserve global dissimilarities; that is, two points that are relatively distant in the data space are guaranteed to be distant in the latent space (Lawrence, 2005; Lawrence and Quiñonero Candela, 2006). Generative topographic mapping, density networks, and GPLVM are examples of these techniques. Among these, GPLVM is a kernel-based method that preserves global dissimilarity. The kernel in kernel-based manifold learning techniques, such as Isomap and GPLVM, captures the nonlinear structures within the data in a non-parametric fashion.

The methodology proposed in this study combines the strengths of both local and global (dis) similarity preservation by utilising GPLVM and Isomap. Specifically, Isomap is used as an initialisation method for GPLVM; we name it–Isomap-GPLVM. Furthermore, since manifold learning is performed to reduce the temporal dimension, the method takes into account the temporal structures present within the EEG data. Consequently, the spatio-temporal local similarities and global dissimilarities within the EEG data are preserved in the latent space. The resulting kernel matrix from GPLVM provides a generic measure of (dis) similarity between EEG channels, capturing the preserved information.

Gaussian Process Latent Variable Model (GPLVM)

As a probabilistic nonlinear manifold learning technique, a GPLVM (Lawrence, 2003) learns the mapping of a high-dimensional observed dataset $\mathbf{Y} \in \mathbb{R}^{N\times D}$ from the corresponding low-dimensional latent positions $\mathbf{X} \in \mathbb{R}^{N\times Q}, Q < D$, i.e. a mapping from $\mathbf{X} \rightarrow \mathbf{Y}$, using a Gaussian process (GP) (Schulz et al., 2018). Here $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_N]^T, \mathbf{y}_i \in \mathbb{R}^D$ and $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N]^T, \mathbf{x}_i \in \mathbb{R}^Q$.

In principal component analysis (PCA), the mapping $\mathbf{X} \rightarrow \mathbf{Y}$ is governed by the dominant eigenvectors of the covariance matrix (Tipping and Bishop, 1999). GPLVM is a probabilistic manifold learning method, which is a nonlinear generalisation of PCA (Lawrence, 2003), where the probabilistic mapping $\mathbf{X} \rightarrow \mathbf{Y}$ is governed by a kernel matrix $\mathbf{K} \in \mathbb{R}^{N\times N}$ (Lawrence, 2003). The marginal log-likelihood of the data $\mathbf{Y}$ given the latent positions $\mathbf{X}$ (Lawrence, 2003; Lawrence, 2005) is

$$L = -\frac{DN}{2}\ln(2\pi) - \frac{D}{2}\ln(|\mathbf{K}|) - \frac{1}{2}\mathrm{tr}(\mathbf{K}^{-1}\mathbf{Y}\mathbf{Y}^T), \quad (1)$$

where $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is a positive semi-definite matrix. The i^{th} row and j^{th} column of $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is given by $k(\mathbf{x}_i, \mathbf{x}_j)$ where $k(\cdot, \cdot)$ is the kernel/covariance function with a set of hyper-parameters $\boldsymbol{\theta}$. The use of a kernel function allows the nonlinear functional mapping from $\mathbf{X}$ to $\mathbf{Y}$ and provides a probabilistic nonlinear latent variable model (Lawrence, 2005). In GPLVM, maximising L is done with respect to both $\mathbf{X}$ and $\boldsymbol{\theta}$, therefore, the optimal estimates for $\mathbf{X}$ and $\boldsymbol{\theta}$ are obtained jointly. This is a highly complex optimisation with the possibility of multiple local minima (Lawrence, 2003). As such, an appropriate initialisation of the latent positions, $\mathbf{X}$, is critical to guide the optimisation of GPLVM (Bitzer and Williams, 2010). Which initialisation method to use depends on the specific application (Lawrence, 2005).

In this study, our objective is to learn the spatial and temporal structures within the EEG data, taking into account both local similarities and global dissimilarities. Since GPLVM only preserves global dissimilarities (Lawrence and Quiñonero Candela, 2006), initialising $\mathbf{X}$ with respect to local similarities within the data is appropriate (Lawrence and Quiñonero Candela, 2006; Lawrence, 2005; Bitzer and Williams, 2010). Isomap has previously been successfully applied in spatio-temporal motion capture data to build latent spaces for controlling a robotic hand (Tsoli and Jenkins, 2008). Furthermore, the use of Isomap to initialise GPLVM has been reported to have superior performance in motion capture data (Bitzer and Williams, 2010). Therefore, Isomap is deemed an appropriate method to determine the initial latent positions. Section 'Isomap as an initialisation for GPLVM' provides further information about the specific Isomap variant used and its superiority over other methods that learn local similarities.

Covariance function in GPLVM. In a GP, the covariance function $k(\cdot, \cdot)$ determines what type of

functions can be learned (Abdessalem et al., 2017). Furthermore, it is the covariance function that defines the regions of similarity and dissimilarity between the input variables (Rasmussen and Williams, 2006). Therefore, in GPLVM, $k(\cdot,\cdot)$ defines the regions of similarity and dissimilarity between the latent positions $\mathbf{x}_i \in \mathbf{X}$.

In this study, the Radial Basis Function (RBF) (also called the squared exponential kernel) (Rasmussen and Williams, 2006) is used as the covariance function for GPLVM. This is due to its inherent properties and the ability to clearly interpret its hyper-parameters (Abdessalem et al., 2017). The RBF kernel has the universal approximating property (Micchelli et al., 2006) and can be integrated against most functions to obtain a smooth mapping from $\mathbf{X} \rightarrow \mathbf{Y}$ (Abdessalem et al., 2017; Rasmussen and Williams, 2006). The RBF covariance function is given by

$$k(\mathbf{x}_i, \mathbf{x}_j) = \sigma^2 \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2l^2}\right), \quad (2)$$

where l and σ are the length-scale and the output-variance hyper-parameters, $\theta = [l, \sigma]$. Here, the length-scale l determines how quickly the similarity between $\mathbf{x}_i$ and $\mathbf{x}_j$ drops off as the distance between the latent positions increases (Abdessalem et al., 2017; Rasmussen and Williams, 2006).

Isomap as an initialisation for GPLVM

Isomap (Tenenbaum et al., 2000) aims to preserve the geometry within nonlinear data by using the geodesic distances (along the surface of the high dimensional data) between the data points. It approximates the geodesic distances using weighted neighbourhood graphs to project high-dimensional data to a lower-dimensional representation, preserving shape information (Tenenbaum et al., 2000). This is the reason for choosing Isomap over methods such as KPCA and t-SNE as the initialisation for GPLVM. The robust kernel Isomap variant (Choi and Choi, 2007) is used in this study.

Robust kernel Isomap approximates the geodesic distance to project the data into the latent space (i.e. lower-dimensional representation) while preserving topological stability and providing a method for eliminating critical outliers (Choi and Choi, 2007). The data points are projected, according to how close the points are, in the data space (i.e. preserving local similarities). In the analysis of EEG data, robustness to noise is vital as this could affect the local similarities and the geodesic distance calculations. This is the main reason for utilising robust kernel Isomap, instead of the competing LLE method and its variants.

Kernel-Based nonlinear Manifold Learning of High-dimensional EEG Data Using Isomap-GPLVM

Isomap-GPLVM is applied individually to the EEG data of each AD and HC participant by reducing the temporal dimension. Following the definition of the data space $\mathbf{Y} \in \mathbb{R}^{N \times D}$ (in Section 'Gaussian Process Latent Variable Model (GPLVM)'), here $N = 23$ (23 EEG channels, Section 'EEG data') and and D is the temporal dimension to be reduced. The associated latent space of each AD and HC participant will be $\mathbf{X} \in \mathbb{R}^{N \times Q}$. The resulting kernel matrix, $\mathbf{K}(\mathbf{X}, \mathbf{X}) \in \mathbb{R}^{23 \times 23}$ of each participant, quantifies the spatio-temporal (dis) similarity information between the 23 respective EEG channels as a generic measure of similarity. Fig. 1A illustrates this.

Robust Kernel Isomap is a technique that approximates the geodesic distance between data points to project them onto a lower-dimensional representation (latent space), while preserving local similarities (how close data points are) (Choi and Choi, 2007; Lawrence and Quiñonero Candela, 2006). The size of the resulting lower-dimensional representation, denoted as Q, is determined by the user. In our approach, we first apply Robust Kernel Isomap to the high-dimensional data to obtain an initial estimate of the lower-dimensional representation. We then use GPLVM to refine the lower-dimensional representation based on the global dissimilarities between data points (how far apart data points are). This will result in the final latent-space $\mathbf{X}$. The kernel matrix, $\mathbf{K}(\mathbf{X}, \mathbf{X})$, from GPLVM governs the mapping between the latent space $\mathbf{X}$ and the high-dimensional data space $\mathbf{Y}$ (see Section 'Gaussian Process Latent Variable Model (GPLVM)'). This kernel matrix $\mathbf{K}$ will reflect both local similarities and global dissimilarities that are learnt from the original data and embedded in $\mathbf{X}$, and we refer to it as the kernel (dis) similarity matrix. The use of the RBF covariance function allows to quantify this information in $\mathbf{K}$ as a generic measure of similarity. By reducing the temporal dimension of the data, temporal information is naturally incorporated into the kernel (dis) similarity matrix.

Given the choice of the latent dimension, Q, the best set of values, from a grid search, for $\theta = [l, \sigma]$ is chosen based on how well the kernel (dis) similarity matrices are distinguishable from HC to AD. Linear SVM with Monte-Carlo cross validation (SVM-MCV) is used to assess this (see Section 'Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure' and Fig. 1B). The initial conditions that result in the highest average area under the receiver operator curve (AU-ROC) are chosen. The grid search is done using the search ranges $l = [2, 100]$ and $\sigma = [2, 30]$ for several latent dimensions, $Q = [5, 8]$.

It was found in this study, that fixing the length-scale l of the RBF covariance function $k(\cdot,\cdot)$ in eq. (2), in the maximising of L in eq. (1), produces consistently better average AU-ROC results from applying SVM-MCV (Fig. 1B and Section 'Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure') across latent dimensions $Q = [5, 8]$. Similar behaviour has been reported in (Rastgoufard and Alsamman, 2016) when using GPLVM with a back-constrained (Lawrence and Quiñonero Candela, 2006) likelihood to preserve local similarities. The appropriate fixed value of l and the initial condition of σ in $k(\cdot,\cdot)$, in maximising L, is found using a grid search method. Therefore, in other words, the fixed value for the length-scale l, as mentioned above, leads the manifold learning method to produce a kernel matrix where its similarity measure is optimised for the differen-

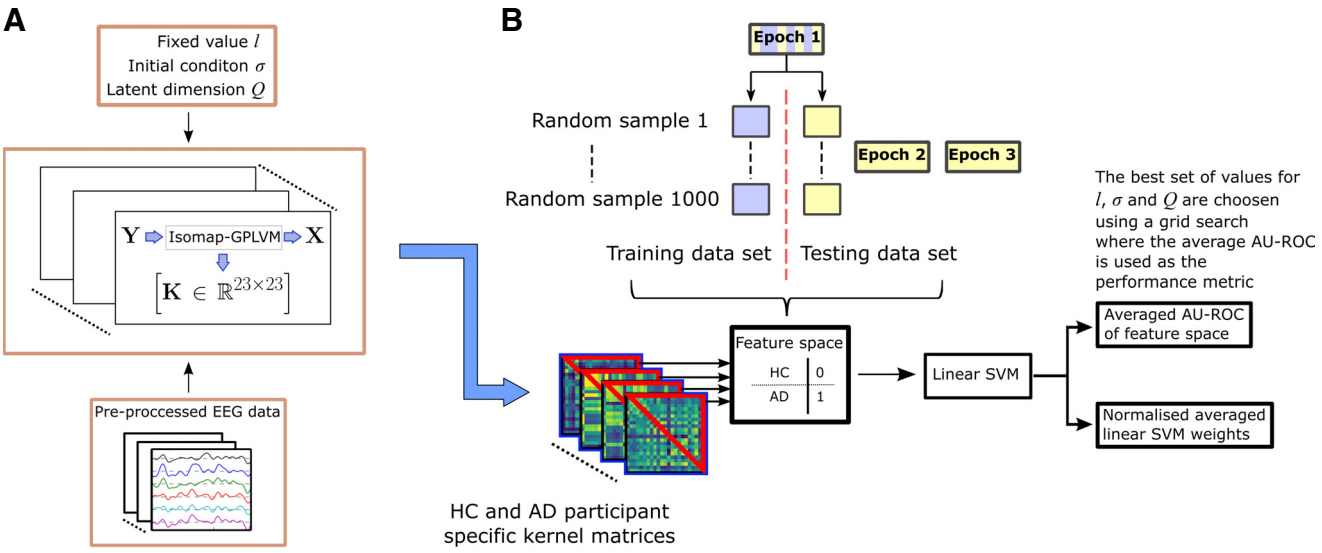


Fig. 1. The Isomap-GPLVM method for evaluating the kernel (dis) similarity matrices. **A) Isomap-GPLVM**: EEG data of each participant is first pre-processed via FFT filtering, to remove unwanted frequency components and normalise the data (zero mean and unit variance). Then participant specific kernel (dis) similarity matrices are evaluated using Isomap-GPLVM. From the EEG data, **Y**, Isomap-GPLVM learns the spatio-temporal local similarities and global dissimilarities within the data (see Section 'Kernel-Based nonlinear Manifold Learning of High-dimensional EEG Data Using Isomap-GPLVM'). This information is embedded in the latent space **X** and is reflected in the kernel matrix **K** (see Section 'Gaussian Process Latent Variable Model (GPLVM)'). The best set of values for l, σ and Q, from a grid search, are chosen based on how well the kernel (dis) similarity matrices are distinguishable from HC to AD. **B) Linear SVM-MCV** is used to assess this. The set of values for l, σ and Q that produce the best average AU-ROC from the testing set is chosen. All channel pairwise similarities from the kernel (dis) similarity matrices are used as features. This study uses EEG data from 20 HC and 20 AD participants. Three epochs of EEG data from each participant are available (see Section 'EEG data'). From Epoch 1, 10 HC and 10 AD participants are chosen randomly for the training set, and the remaining 10 HC and 10 AD are used for the testing set. Epochs 2 and 3 are used for the testing set. The feature space has two classes, AD and HC. The classification is binary – AD is denoted as 1 and HC as 0. A linear SVM classifier is used on the feature space to determine which channel pairs (inter-relationships) are better at distinguishing. between groups.

tiation of HC and mild to moderate AD EEG data. The complete Isomap-GPLVM methodology is illustrated in Fig. 1.

It should be noted, that the pre-processed 23-channel EEG data of each participant contains 2400 time samples (see Section 'EEG data'), $\mathbf{Y} \in \mathbb{R}^{23\times2400}$. This can be nearly perfectly represented (recovered from $\mathbf{X} \rightarrow \mathbf{Y}$ with a 95% confidence) in a latent space $\mathbf{X} \in \mathbb{R}^{23\times Q}$, with $Q \geqslant 5$ using Isomap-GPLVM. To achieve the same recovery accuracy, the linear principal component analysis, requires a latent dimension of $Q = 20$.

Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure

This study comprises of 20 HC and 20 AD participants (see Section 'EEG data'). From each participant, three 12-s epochs of EO EEGs are used. The kernel (dis) similarity matrices of the EEG data are produced for each AD and HC participant using Isomap-GPLVM, for all three epochs. The pairwise (dis) similarity measures are used as features, to assess how well it can distinguish between the HC and AD groups.

Due to the relatively small number of participants, linear SVM is preferred, as it has been shown to be effective with small datasets (Moctezuma and Molinas, 2020; Lotte et al., 2018; Zhang et al., 2017). Furthermore, it provides a globally optimum solution and the number of features does not affect the classification complexity (Moctezuma and Molinas, 2020; Lotte et al., 2018; Joachims, 1999). The use of a Monte-Carlo cross-validation strategy is preferred, because of its better performance with smaller data samples (Shan, 2022) and the asymptotically consistent property for linear (classification) models (Shao, 1993). Additionally, some AD participants could easily be detected, while others might not. This depends on the severity of neurodegeneration of the participants with mild to moderate AD used in this study. Since such information is not explicitly available, a randomised cross-validation strategy is used to obtain a fair balance in the linear SVM weights (Shao, 1993; Xu and Liang, 2001; Xu et al., 2004). Furthermore, this also implies that, including features from more than one epoch of each participant in the training dataset, could increase the risk of participant-specific biases in the classifier. In this study, we aim to find the group differences in FC, between HC and AD and thereby to perform channel selection. Therefore, the use of only one epoch in the training dataset and the rest in the testing set, ensures the generalising capability of the linear SVM classifier and the weightings skewed towards the most significant predictors.

Monte Carlo cross-validation is used where, from the *first epoch*, 10 HC and 10 AD participants are randomly picked for the training set. The remaining 10 HC and 10 AD participants from the *first epoch* are used for testing. 1000 such random samples are taken to generate 1000 different training and testing sets. As discussed previously, the 2nd and 3rd epochs of all participants will also be included in the testing dataset (Fig. 1B)).

AU-ROC from the 1000 *testing sets* are used as a metric to determine the performance of the linear SVM classification. This procedure of linear SVM-MCV is illustrated in Fig. 1B. The AU-ROC is preferred when considering the cost of misclassification, especially in medical diagnosis, as it helps to minimise the likelihood of misdiagnosis (Zweig and Campbell, 1993; Hossin and Sulaiman, 2015; Carter et al., 2016).

Kernel (dis) similarity matrix analysis

After the initial condition that produces the highest average AU-ROC is determined, SVM-MCV is used to analyse the associated kernel (dis) similarity matrices and rank the pairwise channel FC changes between HC and AD. The ranking is done using the absolute values of the normalised average of the linear SVM weights (normalised average linear SVM weights) resulting from the 1000 training sets (Section 'Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure'). The averaged weights are normalised, so that, the highest absolute weight is 1 (Fig. 1B). Due to the linearity in the classification method used, according to the superposition principle, the averaging of the linear SVM weights can be easily interpreted. Two approaches are used, when implementing SVM-MCV:

A). **Global EEG FC analysis**. All pairwise (dis) similarities are used, as in Fig. 1B. This identifies the best channel pairwise comparisons that can distinguish between HC and AD considering the global EEG interactions.

B). **Channel-specific EEG FC analysis**. Each row of all kernel matrices forms a channel-specific feature space, as shown in Fig. 2. SVM-MCV is applied to each feature space individually. This ranks channel pairs, considering a specific channel and its connectivity with the rest of the EEG to identify any significant region-specific FC changes between the HC and AD groups (Neufang et al., 2011; Babiloni et al., 2004; Babiloni et al., 2006).

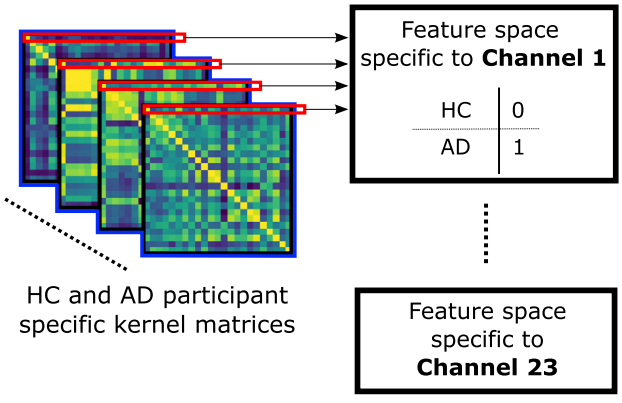


Fig. 2. SVM-MCV: Channel-specific EEG FC analysis. All corresponding rows from HC and AD kernel matrices are grouped into channel-specific feature spaces. Each feature space has two classes, i.e. AD (1) and HC (0). Individual linear SVM classifiers are used on each feature space to determine which EEG channels, considering only its connectivity with the rest of the EEG, are better at distinguishing between the groups. The same SVM-MCV approach (as described in Section 'Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure' and Fig. 1B), is now applied to each individual channel-specific feature space.

Software packages used

The Isomap-GPLVM methodology is implemented in Python. GPLVM is applied using the package 'GPflow' (Matthews et al., 2017). Robust kernel Isomap and the various distance measures (Euclidean, Bray-Curtis, Correlation) are implemented using, the 'Scikit-learn' package (Pedregosa et al., 2011). The 'Dyconnmap' package (Marimpis et al., 2021), is used for the FC measures PLV, iPLV, PLI and iCoherence. Multiple hypothesis testing (in Section 'Results') is done using the Mann–Whitney U test, using the 'SciPy' package (Virtanen et al., 2020), and the Benjamini-Hochberg (Benjamini and Hochberg, 1995) false discovery rate controlling method, using the 'MultiPy' package (Puoliväli et al., 2020).

RESULTS

The Isomap-GPLVM method introduced in Section 'Experimental procedures' is applied to the EO and EC EEG data (Section 'EEG data') separately. The best fixed value l and the initial condition σ for $k(\cdot,\cdot)$ to maximise L, is determined via a grid search using the three 12-s epochs according to the procedure explained in Section 'Kernel-Based nonlinear Manifold Learning of High-dimensional EEG Data Using Isomap-GPLVM', using SVM-MCV (see Section 'Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure' and Fig. 1B).

Table 1 illustrates, the selected Q and l values for EO and EC conditions. Given the choices for l, σ and Q, from the participant-specific kernel (dis) similarity matrices evaluated (Fig. 1 A), the channel inter-relationships (FC) that are able to differentiate well, between AD and HC groups are presented in this section, for both EO and EC conditions. The FC analysis is done in two approaches: global EEG FC changes and channel-specific EEG FC changes (see Section 'Kernel (dis) similarity matrix analysis').

Fig. 3 illustrates the bipolar montage EEG channels used in this work (see Section 'Data') on a 10–20 international standard electrode placement map. The mid-points between the 10–20 EEG overlap with certain 10–10 EEG electrode positions (Jurcak et al., 2007). Therefore, the EEG channels used in this work (Fig. 3) measure the scalp electrical activity at those overlapping positions (see Section 'Data'). The corresponding under-

Table 1. Selected latent dimension and fixed length-scale values for EO and EC conditions.

Condition	Latent dimension Q	Fixed length-scale l	Average AU-ROC
EO	8	66.5	0.73
EC	8	83.5	0.77

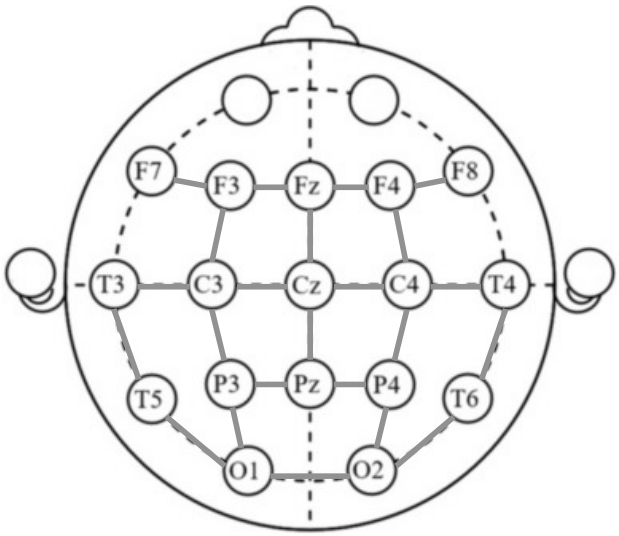


Fig. 3. All the 23 channels, bipolar montage. EEG bipolar montage channels mapped into a 10–20 international standard arrangement. The bold grey lines connecting any two EEG electrodes indicate that these two electrodes result in a bipolar channel. Bipolar channels give an estimate of the instantaneous electric field along the scalp surface midway between the pair of electrodes.

lying cortical regions of these positions (Rojas et al., 2018) are used as location markers. Table 2 shows the 23 bipolar montage EEG channels used and the respective underlying cortical regions.

It should be noted, that the EEG has a low spatial resolution. EEG bipolar channels measure the propagated electrical activity on the overlying scalp regions (Table 2). Therefore, in this study when results are presented with respect to the cortical region, it does not refer to the explicit activity in the actual brain cerebral cortex.

Table 2. List of all 23 channels of the scalp EEG bipolar montage and the corresponding underlying cortical regions.

Channel index and name		Corresponding cortical region
0	O1-O2	Occipital (O)
1	P4-O2	
2	P3-O1	
3	T5-O1	
4	T6-O2	
5	P3-PZ	Parietal (P)
6	P4-PZ	
7	T3-T5	Temporal (T)
8	T4-T6	
9	C3-P3	Centro-Parietal (CP)
10	C4-P4	
11	CZ-PZ	
12	C3-CZ	
13	C4-CZ	
14	T3-C3	Centro-Temporal (CT)
15	T4-C4	
16	F3-C3	Fronto-Central (FrC)
17	FZ-CZ	
18	F3-FZ	
19	F4-FZ	
20	F4-C4	
21	F7-F3	Frontal (F)
22	F8-F4	

Comparison of kernel (dis) similarity against commonly used functional connectivity measures

To demonstrate the efficacy of the proposed FC measure, comparisons with commonly used FC measures are presented here. Table 3 shows the comparison of the AU-ROC values for the SVM-MCV global EEG FC analysis, for both EO and EC conditions. It is evident that, when considering all the pairwise kernel (dis) similarity measures (global EEG FC analysis), our proposed Isomap-GPLVM based FC measure, has a considerably higher AU-ROC than other measures. This is especially true for the EC condition.

Fig. 4 illustrates the performance of our Isomap-GPLVM-based FC measure with respect to the channel-specific EEG FC analysis (see Section 'Kernel (dis) similarity matrix analysis'). The distribution of the averaged AU-ROCs from all channel-specific feature spaces (Fig. 2) of the respective FC measures, is shown as a box-plot in Fig. 4. The channel-specific EEG FC analysis is used to identify the important FC changes, between HC and AD, with respect to a specific cortical region. In both the EO and the EC conditions, the average AU-ROC of each channel-specific feature space is directly compared, between our method and other FC measures. In both conditions, it was observed that half of the channel-specific feature spaces (Fig. 2) from our method attained higher AU-ROC values, than the corresponding feature spaces in other FC measures. From the remaining half, most feature spaces matched the performance of corresponding feature spaces in other FC measures, while the rest underperformed. The data for all the average AU-ROCs of all channel-specific feature spaces for all FC measures used, in this comparison is not provided here. However, what is mentioned above is reflected in the box-plots in Fig. 4. Therefore, in general, with respect to the channel-specific EEG FC analysis, the proposed method improves the overall result under both conditions. In the EO condition, our method performs significantly better compared to other FC measures.

Table 3. Comparison of SVM-MCV global analysis with AU-ROC values from the proposed FC measure against commonly used FC measures, under EO and EC conditions.

FC measure	AU-ROC EC condition	AU-ROC EO condition
Isomap-GPLVM	**0.77 ± 0.07**	**0.73 ± 0.03**
Euclidean	0.61 ± 0.04	0.62 ± 0.04
Bray-Curtis	0.60 ± 0.02	0.61 ± 0.03
Correlation	0.68 ± 0.05	0.67 ± 0.04
PLV	**0.74 ± 0.04**	**0.70 ± 0.04**
iPLV	0.59 ± 0.05	0.64 ± 0.05
PLI	0.58 ± 0.06	0.63 ± 0.05
iCoherence	0.62 ± 0.04	0.58 ± 0.04

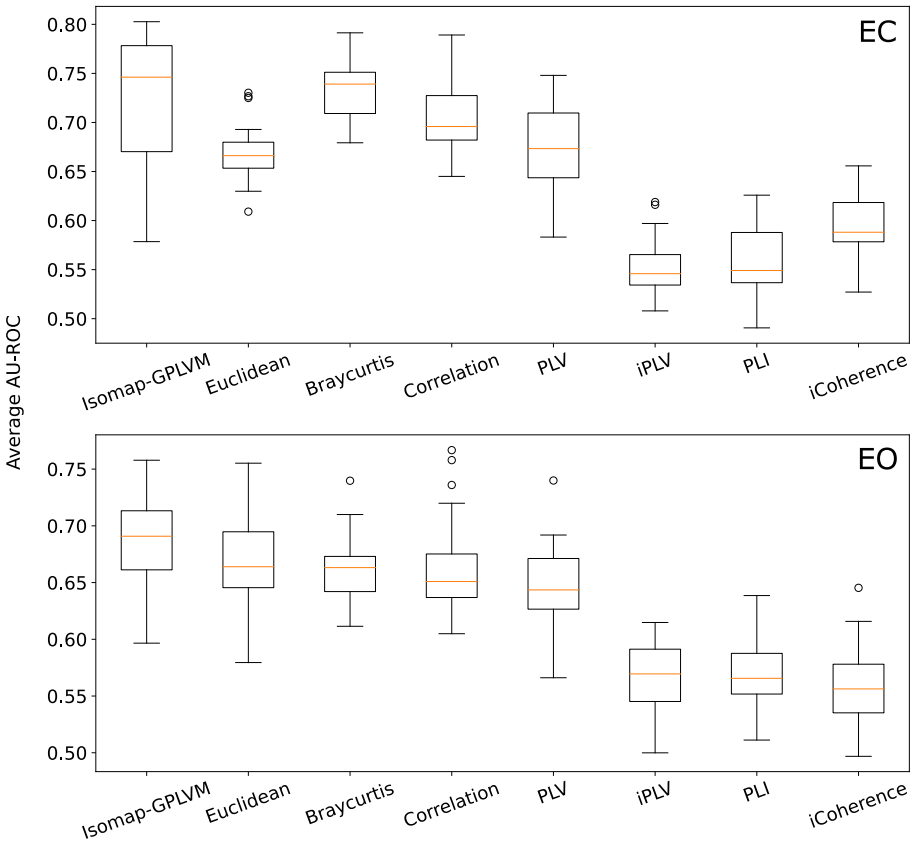


Fig. 4. Comparison of the proposed Isomap-GPLVM FC measure against commonly used measures using SVM-MCV channel-specific approach. The distribution of the average AU-ROC across all channel-specific feature spaces in each FC measure is shown for both EO and EC conditions.

Kernel (dis) similarity matrices of HC and AD groups

The Mann–Whitney U test (Mann and Whitney, 1947) is used for the element-wise statistical comparison of the kernel matrices between the HC and AD groups. Due to the multiple statistical comparisons done here, the *p*-values need to be approximately corrected (Dauwels et al., 2010b; Puoliväli et al., 2020). Also, due to the large number of comparisons (i.e. 23 EEG channels correspond to 253 channel combinations), controlling the false discovery rate (i.e. positive results that could be in fact negative) (Benjamini and Hochberg, 1995)) is preferred over controlling the family-wise error rate (Dauwels et al., 2010b; Puoliväli et al., 2020). Therefore, the Benjamini-Hochberg (Benjamini and Hochberg, 1995) false discovery rate controlling (FDR) method is used to obtain the corrected *p*-values. Pairwise kernel (dis) similarities (FC measures) that have statistically significant differences between the HC and AD groups (*p*-values < 0.05) are denoted as–1's in significance matrix $\mathbf{S} \in \mathbb{R}^{23\times23}$, zero otherwise. The significance matrices for both the EC and EO conditions are illustrated in Fig. 5, where blue elements indicate, the statistically significant changes in the pairwise connectivities between the HC and AD groups.

From Fig. 5, it is evident, that there are localised FC changes within certain underlying cortical regions (e.g. within the centro-parietal EEG region) and global EEG FC changes between regions (e.g. between the centro-parietal and occipital EEG regions). This can be a reflection of the specific patterns of dysfunction that have been mentioned in the literature, in which AD EEG data exhibits a specific change in FC compared to HC (Stam, 2005; Dauwels et al., 2010b; Dauwels et al., 2011) and connectivity in certain regions of the EEG being affected (Falk et al., 2012; Deng et al., 2017; Tylová et al., 2018; Abásolo et al., 2009; Fraga et al., 2013; Al-Qazzaz et al., 2014). These FC changes could be linked to within-frequency and cross-frequency coupling between brain regions (Babiloni et al., 2016; Sadaghiani et al., 2022).

Global functional connectivity changes and channel pair selections

All pairwise kernel (dis) similarities are used (Fig. 1B) to determine, in a global sense, which spatio-temporal FC differences between cortical regions (EEG channels) are more important in distinguishing between the HC and AD groups. Table 4 and 5 shows the top 20 channel pairs that are ranked according to the averaged normalise linear SVM weights (Section 'Kernel (dis) similarity matrix analysis', Fig. 1B) for EO and EC conditions, respectively. EEG channel pairs in these two tables are arranged in a way so that channels related to the same underlying cortical regions can be grouped.

Tables 4 and 5 show that inter-regional FC between EEG channels from the occipital region and other regions, i.e. parietal (P–O, Fig. 6C), centro-parietal (CP–O, Fig. 6A), and fronto-central (FrC–O, Fig. 6B), attain a considerable space among the top 20 rankings. This is observed in both EO and EC conditions. However, in the EC condition this is specific to the right occipital region (channels P4-O2 and T6-O2, Table 5). In particular to the EO condition, as shown in Table 4, FC between EEG channels from the frontal and occipital regions (F–O, Fig. 6D) have a significant presence within the top 20 weightings. Therefore, this suggests that connectivity between the occipital region and those regions mentioned above, respective to each condition, can be important in identifying people with mild to moderate AD as shown in Fig. 6.

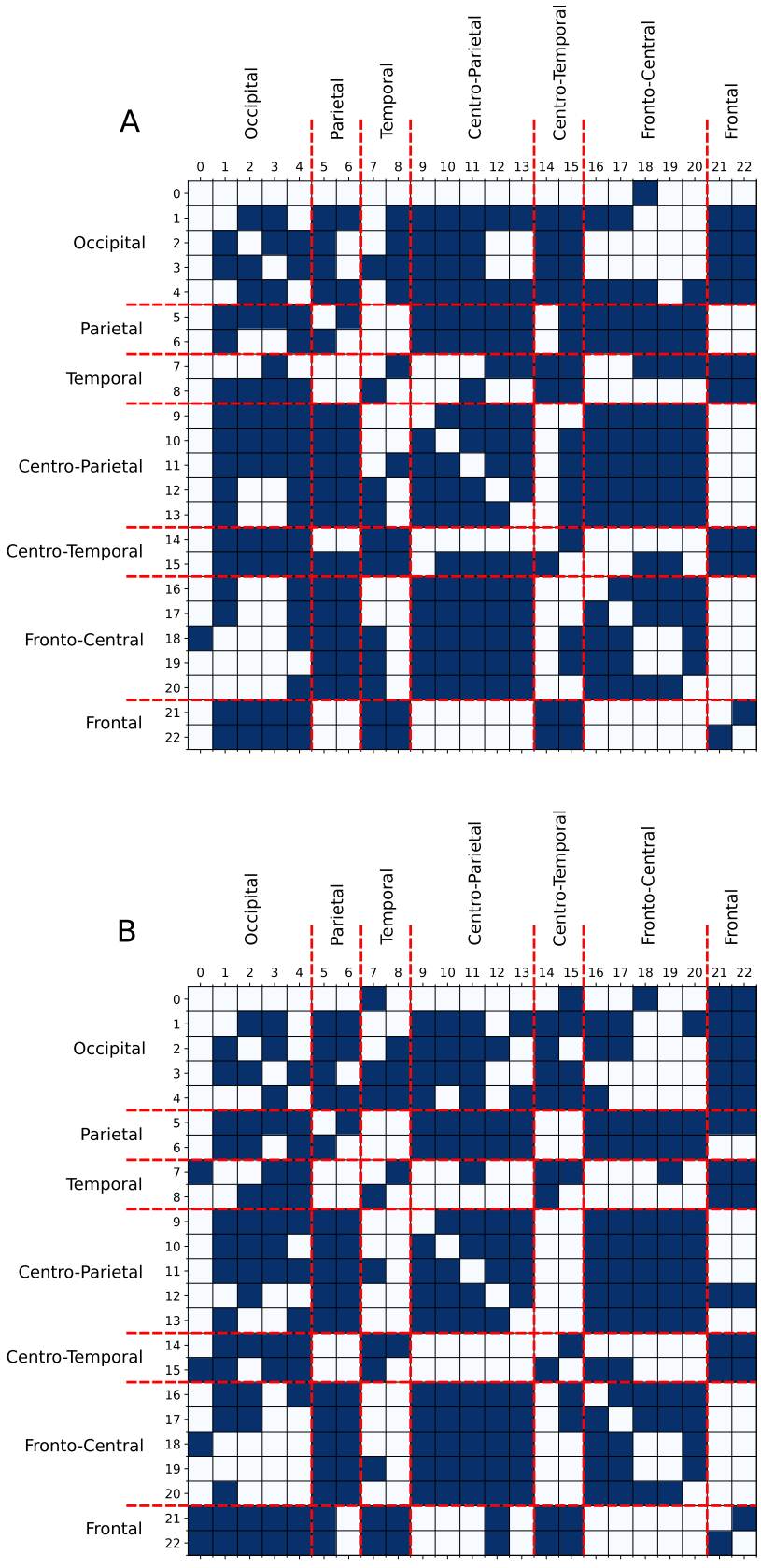


Fig. 5. The significance matrices, S, for both EC (A) and EO (B) conditions. The figures show, for both EC and EO conditions, the statistical significance of all the elements of the kernel (dis) similarity matrices between the HC and AD groups (based on all epochs of all participants). The corresponding channels are provided in Table 2. The significant FC changes are indicated in blue.

Channel-specific functional connectivity changes and channel pair selections

In this section, the results of the channel-specific EEG FC analysis (Section 'Kernel (dis) similarity matrix analysis', Fig. 2), using kernel (dis) similarity matrices, are presented. This approach determines, at the EEG sensor level, significant changes in FC specific to the cortical region between the HC and AD groups. The channel-specific approach provides another layer of information.

To form a channel-specific feature space, each row of all kernel matrices that correspond to a particular channel is used (Fig. 2)). SVM-MCV is then applied to each feature space individually to identify individual EEG channels that exhibit distinguishable changes in FC with the rest of the EEG data. The average AU-ROC of the channel-specific feature space is used as an evaluation metric. The normalised average linear SVM weights (Section 'Linear SVM and Monte-Carlo cross-validation (SVM-MCV) procedure') of the channel-specific feature space are used to rank the importance of FC changes relative to the channel being considered (Section 'Kernel (dis) similarity matrix analysis').

Tables 6 and 7 report channel-specific feature spaces with an average AU-ROC > 0.7 for the EO and EC conditions, respectively. These tables also report the kernel (dis) similarity features that attain a high rank in the feature space being considered (i.e. normalised average linear SVM weight 0.9–1). Fig. 7 illustrates the channel-specific feature spaces with an average AU-ROC > 0.7 mapped to the placement of 10–20 international electrodes.

As seen from Fig. 7, the EEG channels associated with the channel-specific feature spaces with an average AU-ROC > 0.7 lie mostly within the fronto-parietal regions of the cortex for both conditions. However, in the EC condition (Fig. 7A), channel-specific feature spaces associated with the EEG channels in the right

Table 4. EO condition. Ranking of (dis) similarity features of channel-pairs–only the top 20 are shown.

Channel Pairs (Indexes and names)				Averaged normalise linear SVM weight	Ranking	Connecting regions
11	1	CZ-PZ	P4-O2	0.84	6	CP - O
13	1	C4-CZ	P4-O2	0.78	7	
9	1	C3-P3	P4-O2	0.68	12	
12	3	C3-CZ	T5-O1	0.68	13	
13	4	C4-CZ	T6-O2	0.60	19	
21	3	F7-F3	T5-O1	0.88	3	F - O
22	3	F8-F4	T5-O1	0.72	8	
22	4	F8-F4	T6-O2	0.59	20	
18	9	F3-FZ	C3-P3	0.68	11	FrC - CP
18	17	F3-FZ	FZ-CZ	0.68	14	FrC - FrC
16	1	F3-C3	P4-O2	1.00	1	FrC - O
17	1	FZ-CZ	P4-O2	0.88	4	
16	4	F3-C3	T6-O2	0.71	10	
16	2	F3-C3	P3-O1	0.64	16	
20	1	F4-C4	P4-O2	0.63	17	
5	3	P3-PZ	T5-O1	0.91	2	P - O
5	2	P3-PZ	P3-O1	0.86	5	
5	1	P3-PZ	P4-O2	0.72	9	
6	1	P4-PZ	P4-O2	0.66	15	
6	3	P4-PZ	T5-O1	0.62	18	

Table 5. EC condition. Ranking of (dis) similarity features of channel-pairs–only the top 20 are shown.

Channel Pairs (Indexes and names)				Averaged normalise linear SVM weight	Ranking	Connecting regions
9	4	C3-P3	T6-O2	0.99	3	CP - O
11	4	CZ-PZ	T6-O2	0.96	4	
13	4	C4-CZ	T6-O2	0.90	5	
10	4	C4-P4	T6-O2	0.90	7	
11	1	CZ-PZ	P4-O2	0.80	9	
12	4	C3-CZ	T6-O2	0.77	11	
9	1	C3-P3	P4-O2	0.72	14	
22	15	F8-F4	T4-C4	0.68	17	F - CT
21	4	F7-F3	T6-O2	0.62	19	F - O
17	4	FZ-CZ	T6-O2	1.00	1	FrC - O
16	4	F3-C3	T6-O2	1.00	2	
16	1	F3-C3	P4-O2	0.80	10	
17	1	FZ-CZ	P4-O2	0.75	12	
18	4	F3-FZ	T6-O2	0.70	15	
20	4	F4-C4	T6-O2	0.62	20	
1	0	P4-O2	O1-O2	0.67	18	O - O
6	4	P4-PZ	T6-O2	0.90	6	P - O
5	4	P3-PZ	T6-O2	0.86	8	
5	1	P3-PZ	P4-O2	0.74	13	
6	1	P4-PZ	P4-O2	0.70	16	

hemisphere appear to be important (average AU-ROC > 0.7). In the case of mild to moderate AD, significant FC changes between these EEG channels and the rest of the EEG is observed from the proposed FC analysis methodology.

DISCUSSION

The results presented in the previous section indicate that certain key areas of the brain are affected by AD (Figs. 5–7). In order to identify whether our results are consistent with functional Magnetic Resonance Imaging (fMRI) results of mild to moderate AD, first, it is necessary to determine how the bipolar channels used in this study (Section 'EEG data') relate to the functional connectivity networks (Yeo et al., 2011) of the brain.

Yeo *et. al.* (Yeo et al., 2011) revealed the existence of seven primary functional networks using time correlations between the fMRI of 1,075 Regions of Interest (ROI). These networks are shown to be valid for multiple participants and robust against various data processing methods. The seven functional networks are visual network (VN), somatomotor network (SN), dorsal attention network (DAN), ventral attention network (VEN), limbic network (LN), fronto-parietal network (FPN) and default mode network (DMN). Based on (Yeo et al., 2011), Rojas *et. al.* (Rojas et al., 2018) used the electrode positions of the international standard 10–20 EEG and the 10–10

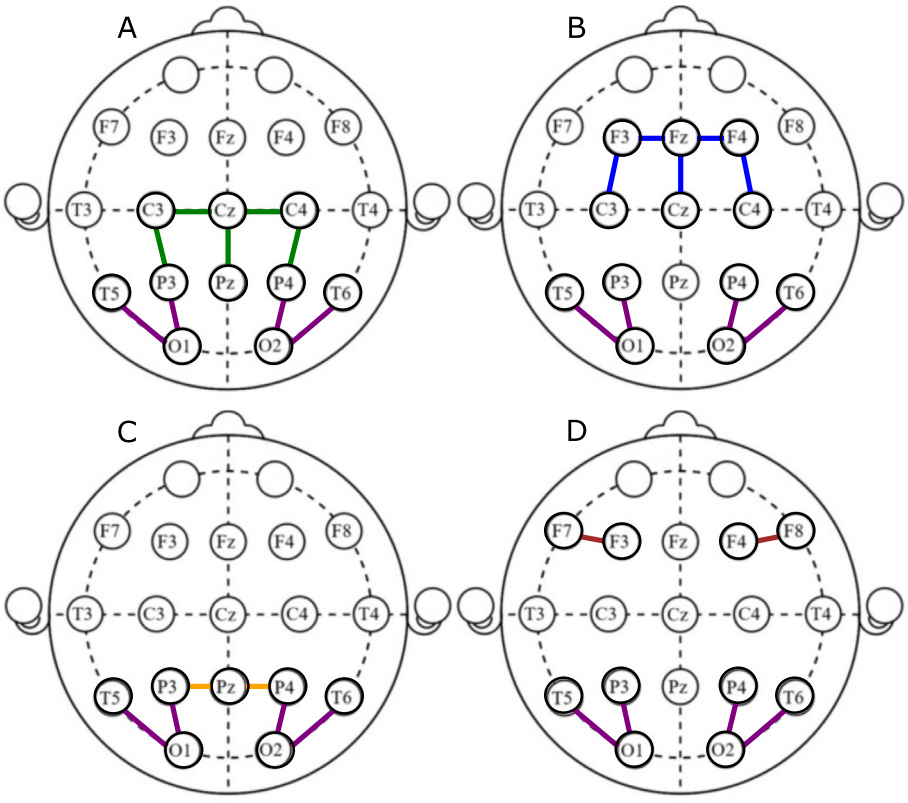


Fig. 6. Inter-regional connectivity between EEG channels from the regions shown can be important in identifying people with mild to moderate AD. **A) CP–O**, **B) FrC–O**, **C) P–O** for both EO and EC conditions while **D) F–O** only for the EO condition (Tables 4 and 5). EEG channels related to the corresponding cortical regions are shown in different colours. Occipital region (O)–purple, Parietal region (P)–orange, Centro-parietal region (CP)–green, Fronto-central region (FrC)–blue and Frontal region (F)–red. It shoud be noted that bipolar channels give an estimate of the instantaneous electric field along the scalp surface midway between the pair of electrodes.

EEG as seed positions to provide a reproducible model demonstrating the relationship between the 10–20 EEG electrode positions and the seven functional networks revealed by (Yeo et al., 2011). This is carried out by simultaneous acquisition of EEG and resting-state fMRI (rs-fMRI). Rojas *et. al.* used the Sørensen–Dice index (F1 score) to quantify the similarities between the positions of the 10–20 electrode placements and the seven functional networks mentioned above. The bipolar EEG channels used in this study estimate the electric field midway between the pair of electrodes that form the said channel (Nunez and Srinivasan, 2006). Therefore, to determine approximate similarities between a bipolar channel (Fig. 3) and and the functional networks, the average of the Sørensen–Dice indices of the two electrodes (Fig. 9 and Supplementary Table 3 in (Rojas et al., 2018)) that form the bipolar channel is used. An example of this is shown in the appendix (Appendix A).

Fig. 8 illustrates the relationship between the bipolar EEG channels used in this study and functional networks FPN, DAN, VAN and DMN. In mild to moderate AD, the connectivity changes within networks FPN, DAN, VAN and DMN have been previously reported to be significant (Babiloni et al., 2016; Neufang et al., 2011; Zhang et al., 2010; Gour et al., 2014). Therefore, our following discussion will only focus on these networks, as shown in Fig. 8.

For the EO condition, changes in FC between the cortical regions (EEG channels) shown in Fig. 6 B and D can be speculated as VAN related (Zhang et al., 2010). Considering both EO and EC conditions, changes in the inter-regional FC (as shown in Fig. 6 A, B, C and D) can be linked to changes in connectivity within the FPN and DMN networks (Zhang et al., 2010). Fig. 6 A and C can

Table 6. The average AU-ROC values of the channel-specific feature spaces for the EO condition. Channel-specific feature spaces with average AU-ROC > 0.7 are only shown.

Channel-specific feature space	Channels with normalised average weight 0.9–1	Average AU-ROC of feature space
F3-FZ	P4-O2	0.758
FZ-CZ	P4-O2, F3-FZ	0.737
F3-C3	P4-O2	0.734
C3-P3	P4-O2, T5-O1	0.733
CZ-PZ	P4-O2, F3-FZ	0.719
F4-C4	P4-O2, F3-FZ	0.714
P3-PZ	T5-O1	0.713
C4-P4	P4-O2	0.712
T5-O1	T6-O2, F7-F3	0.710
C3-CZ	P4-O2, T5-O1	0.706

Table 7. The average AU-ROC values of the channel-specific feature spaces for the EC condition. Channel-specific feature spaces with average AU-ROC > 0.7 are only shown.

Channel-specific feature space	Channels with normalised average weight 0.9–1	Average AU-ROC of feature space
C3-P3	T6-O2	0.803
CZ-PZ	T6-O2	0.799
T6-O2	T5-O1, T3-T5, F7-F3	0.797
C4-P4	T6-O2	0.792
F3-FZ	T6-O2	0.790
P3-PZ	T6-O2	0.785
FZ-CZ	T6-O2	0.772
C3-CZ	T6-O2	0.769
F3-C3	T6-O2	0.764
P4-O2	F3-C3, F4-FZ, F8-F4	0.758
F4-C4	T6-O2	0.757
C4-CZ	T6-O2	0.746
T4-C4	F8-F4	0.721

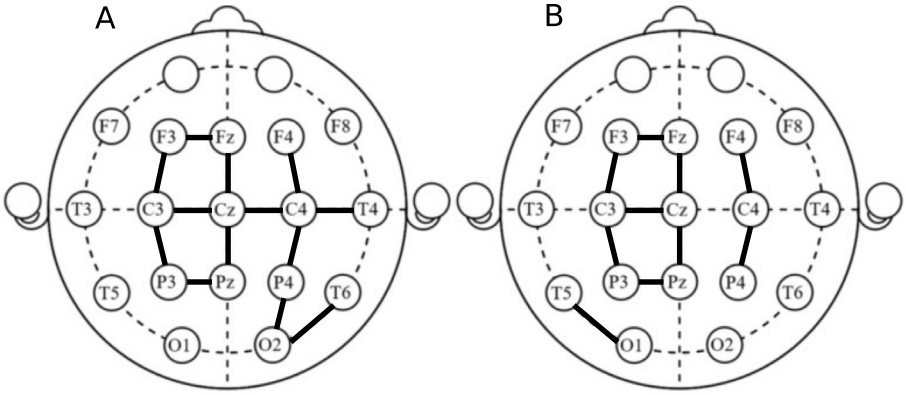


Fig. 7. EC (A) and EO (B) channel-specific feature spaces with average AU-ROC $>$ 0.7. The channel-specific feature spaces with average AU-ROC $>$ 0.7 are mapped into the 10–20 international electrode placement. This is illustrated for both EC and EO conditions. These channels lie mostly within the fronto-parietal regions of the cortex for both conditions.

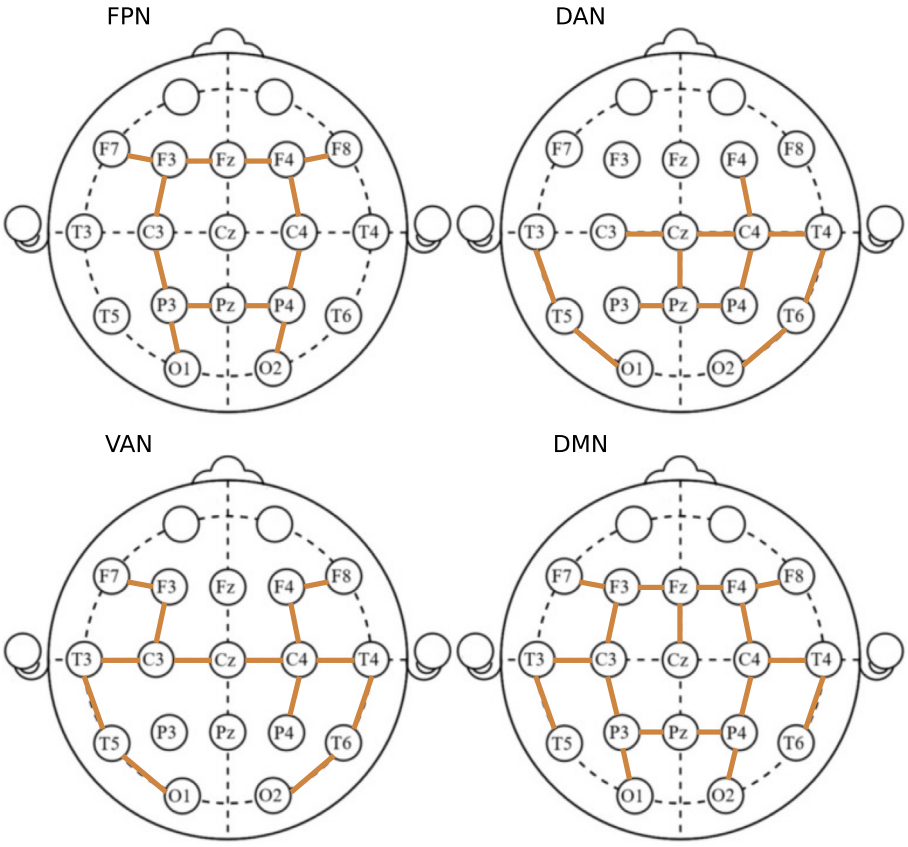


Fig. 8. Relationship between the bipolar EEG channels used and the functional networks. The functional networks discussed in relation to this study and the bipolar channels used is illustrated here. FPN (fronto-parietal network), DAN (dorsal attention network), VAN (ventral attention network) and DMN (default mode network).

be linked to DAN (Zhang et al., 2010; Gour et al., 2014) while Fig. 6 B can be linked to VAN.

The results in Section 'Channel-specific functional connectivity changes and channel pair selections' show the channels that have the most significant FC changes with the rest of the EEG (Fig. 7). This can be a reflection of the FC changes in FPN, DAN, VAN and DMN networks. The EEG channels shown in Fig. 7 are mostly related to the fronto-parietal region of the cortex (fronto-central and centro-parietal regions combined, Table 2). This region has been reported to play an important role in the diagnosis of AD in several studies using fMRI (Neufang et al., 2011) (prodromal AD), rs-fMRI (Zhang et al., 2010; Gour et al., 2014) (mild, moderate and severe AD) and EEG (Babiloni et al., 2016) (mild AD). Neufang et al. (Neufang et al., 2011) pointed out, at the early stages of AD, the volume of regional grey matter is related to the reduction in the effective connectivity (through dynamic causal modelling) in the fronto-parietal region. While Babiloni *et al.* (Babiloni et al., 2006) found that a measure of nonlinear interdependence (via the synchronisation likelihood) is significantly reduced in the fronto-parietal channels of eyes-closed EEG in mild AD patients. These studies are consistent with our results in showing that the connectivity between the EEG channels in the fronto-parietal region (Fig. 7) and and the rest of the brain regions have significantly changed in mild to moderate AD.

A novel FC analysis and channel selection method based on kernel-based nonlinear manifold learning is presented in this work. The FC measure takes both local and global spatio-temporal (dis) similarities between EEG channels into account and ranks the pairwise FC measures that are better at distinguishing HC from patients with neurodegenerative diseases. We demonstrate how a kernel-based (dis) similarity matrix via manifold learning can be used as a measure of spatio-temporal functional connectivity between EEG channels and to determine the important inter-relationships in characterising patients with mild to moderate AD. The methodology presented can determine changes in cortical (EEG channel) inter-relationships that are crucial in distinguishing AD patients from HCs. Furthermore, the results reported in our work are consistent with

other previous studies while linking connectivity changes to functional networks.

The main purpose of this paper is to introduce this novel FC analysis and channel selection methodology and its computational procedure. We also demonstrate its efficacy against other commonly used FC measures. The robustness of our method against volume conduction effects in the EEG could be further assessed (Briels et al., 2020). The Isomap-GPLVM method can be made to explicitly control the compromise between local similarity and global dissimilarity information being learnt. This can be achieved by including appropriate prior probabilities of the latent positions, $p(\mathbf{X})$, in the GPLVM log-likelihood (Urtasun et al., 2008; Buettner and Theis, 2012). This development can be used to improve the classification performance further. With the above considerations, the proposed methodology can be developed into a diagnostic tool not only for the detection of neurodegenerative diseases, but also to determine the important FC changes related to the disease.

An important future study is to investigate the detailed forms of nonlinearity using nonlinear dynamic modelling (Billings, 2013) and nonlinear causality measures in the time and frequency domains (Zhao et al., 2013; He et al., 2014a; He et al., 2014b). These in-depth dynamical analysis methods will be applied to the channel pairs and regions determined using the Isomap-GPLVM method. Thus, enabling the further study of the underlining dynamic processes, linear and nonlinear dynamic features, in patients with Alzheimer's disease.

ETHICAL STATEMENT

The EEG data collection study (seperate from this study), conducted by DJB and PSG, was approved by the Yorkshire and The Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). All participants gave their informed written consent.

ACKNOLWEDGEMENT

RG and FH acknowledge Coventry University for the Trailblazer Ph.D. studentship. The EEG data was funded by a grant from Alzheimer's Research UK, grant reference number ARUK-PPG20114B-25. This is a summary of independent research carried out at the NIHR Sheffield Biomedical Research Centre (Translational Neuroscience). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

Abdessalem AB, Dervilis N, Wagg DJ, Worden K (2017) Automatic kernel selection for gaussian processes regression with approximate bayesian computation and sequential monte carlo. Front Built Environ 3.

Abásolo D, Hornero R, Espino P (2009) Approximate entropy of EEG background activity in Alzheimer's disease patients. Intell Automat Soft Comput 15(4):591–603.

Al-Qazzaz NK, Ali SHBM, Ahmad SA, Chellappan K, Islam MS, Escudero J (2014) Role of EEG as biomarker in the early detection and classification of dementia. Sci World J 2014:906038.

Alotaiby T, El-Samie FEA, Alshebeili SA, Ahmad I (2015) A review of channel selection algorithms for EEG signal processing. EURASIP J Adv Signal Process 2015(1):66.

Babiloni C, Barry RJ, Basar E, Blinowska KJ, Cichocki A, Drinkenburg WH, Klimesch W, Knight RT, Lopes da Silva F, Nunez P, Oostenveld R, Jeong J, Pascual-Marqui R, Valdes-Sosa P, Hallett M (2020) International federation of clinical neurophysiology (ifcn) – eeg research workgroup: Recommendations on frequency and topographic analysis of resting state eeg rhythms. part 1: Applications in clinical research studies. Clin Neurophysiol 131(1):285–307.

Babiloni C, Ferri R, Binetti G, Cassarino A, Forno GD, Ercolani M, Ferreri F, Frisoni GB, Lanuzza B, Miniussi C, Nobili F, Rodriguez G, Rundo F, Stam CJ, Musha T, Vecchio F, Rossini PM (2006) Fronto-parietal coupling of brain rhythms in mild cognitive impairment: A multicentric EEG study. Brain Res Bull 69(1):63–73.

Babiloni C, Ferri R, Moretti DV, Strambi A, Binetti G, Dal Forno G, Ferreri F, Lanuzza B, Bonato C, Nobili F, Rodriguez G, Salinari S, Passero S, Rocchi R, Stam CJ, Rossini PM (2004) Abnormal fronto-parietal coupling of brain rhythms in mild Alzheimer's disease: a multicentric EEG study. Eur J Neurosci 19(9):2583–2590.

Babiloni C, Lizio R, Marzano N, Capotosto P, Soricelli A, Triggiani AI, Cordone S, Gesualdo L, Del Percio C (2016) Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms. Int J Psychophysiol 103:88–102.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol) 57(1):289–300.

Billings SA (2013) Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains. John Wiley & Sons.

Bitzer S, Williams CKI (2010) Kick-starting GPLVM optimization via a connection to metric MDS. In: NIPS 2010 Workshop on Challenges of Data Visualization.

Blackburn D, Zhao Y, Marco MD, Bell S, He F, Wei H-L, Lawrence S, Unwin Z, Blyth M, Angel J, Baster K, Farrow T, Wilkinson I, Billings S, Venneri A, Sarrigiannis P (2018) 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 Sci 8(7):134.

Briels CT, Schoonhoven DN, Stam CJ, de Waal H, Scheltens P, Gouw AA (2020) Reproducibility of EEG functional connectivity in Alzheimer's disease. Alzheimer's Res Therapy 12(1):68.

Buettner F, Theis FJ (2012) A novel approach for resolving differences in single-cell gene expression patterns from zygote to blastocyst. Bioinformatics 28(18):i626–i632.

Carter JV, Pan J, Rai SN, Galandiuk S (2016) Roc-ing along: Evaluation and interpretation of receiver operating characteristic curves. Surgery 159(6):1638–1645.

Choi H, Choi S (2007) Robust kernel isomap. Pattern Recogn 40 (3):853–862.

Dauwels J, Vialatte F, Cichocki A (2010a) Diagnosis of Alzheimer's disease from EEG signals: where are we standing? Curr Alzheimer Res 7(6):487–505.

Dauwels J, Vialatte F, Musha T, Cichocki A (2010b) A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. NeuroImage 49(1):668–693.

Dauwels J, Vialatte F-B, Cichocki A (2011) On the early diagnosis of Alzheimer's disease from EEG signals: A mini-review. In: Wang R, Gu F, editors. Advances in Cognitive Neurodynamics (II). Netherlands, Dordrecht: Springer. p. 709–716.

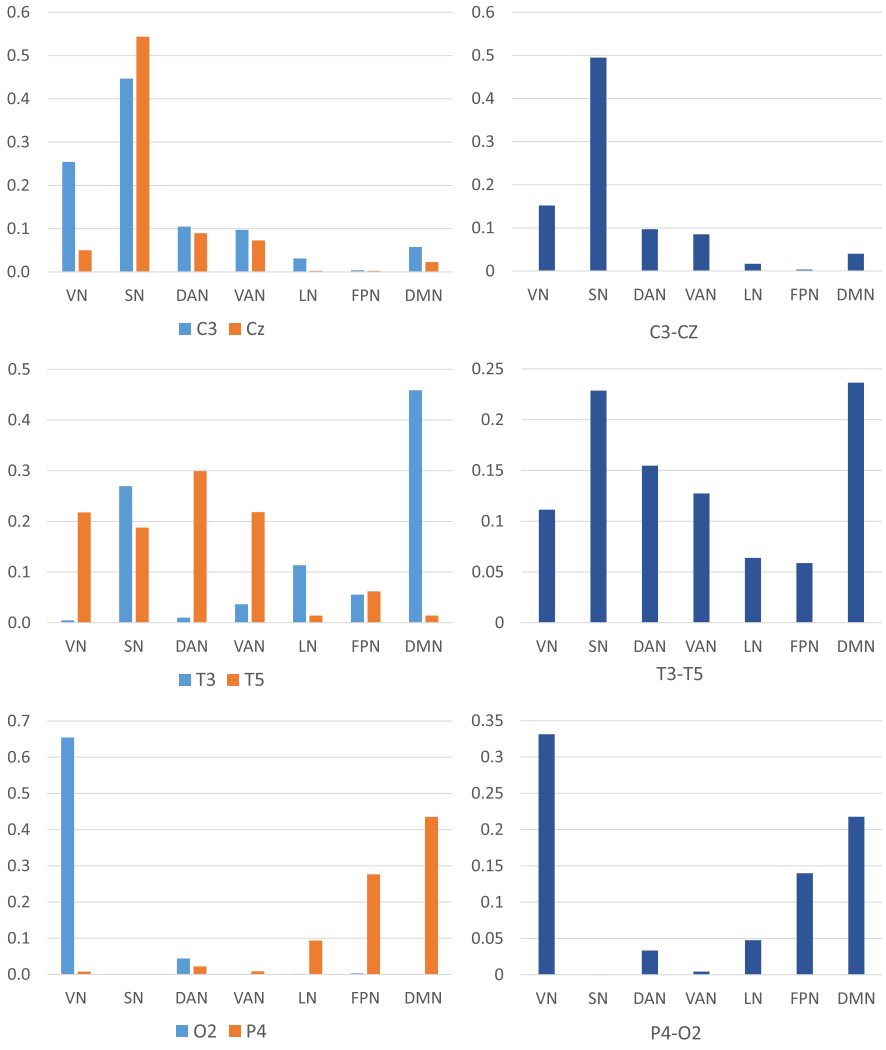
Deng B, Cai L, Li S, Wang R, Yu H, Chen Y, Wang J (2017) Multivariate multi-scale weighted permutation entropy analysis of EEG complexity for Alzheimer's disease. Cogn Neurodyn 11 (3):217–231.

Falk TH, Fraga FJ, Trambaiolli L, Anghinah R (2012) EEG amplitude modulation analysis for semi-automated diagnosis of Alzheimer's disease. EURASIP J Adv Signal Process 2012(1):192.

Fraga F, Falk T, Kanda P, Anghinah R (2013) Characterizing Alzheimer's disease severity via resting-awake EEG amplitude modulation analysis. PloS One 8:e72240.
Gallego-Jutglà E, Solé-Casals J, Vialatte F-B, Dauwels J, Cichocki A (2015) A theta-band eeg based index for early diagnosis of alzheimer's disease. J Alzheimer's Dis 43(1175–1184):4.
Gour N, Felician O, Didic M, Koric L, Gueriot C, Chanoine V, Confort-Gouny S, Guye M, Ceccaldi M, Ranjeva JP (2014) Functional connectivity changes differ in early and late-onset Alzheimer's disease. Hum Brain Mapp 35(7):2978–2994.
He F, Billings SA, Wei H-L, Sarrigiannis PG (2014a) A nonlinear causality measure in the frequency domain: Nonlinear partial directed coherence with applications to EEG. J Neurosci Methods 225:71–80.
He F, Sarrigiannis P, Billings S, Wei H, Rowe J, Romanowski C, Hoggard N, Hadjivassilliou M, Rao D, Grünewald R, Khan A, Yianni J (2016) Nonlinear interactions in the thalamocortical loop in essential tremor: A model-based frequency domain analysis. Neuroscience 324:377–389.
He F, Wei H-L, Billings SA, Sarrigiannis PG (2014b) A nonlinear generalization of spectral granger causality. IEEE Trans Biomed Eng 61(6):1693–1701.
He F, Yang Y (2021) Nonlinear system identification of neural systems from neurophysiological signals. Neuroscience 458:213–228.
Horvath A, Szucs A, Csukly G, Anna Sakovics GS, Kamondi A (2018) Eeg and erp biomarkers of alzheimer's disease: a critical review. FBL 23(2):183–220.
Hossin M, Sulaiman MN (2015) A review on evaluation metrics for data classification evaluations. Int J Data Min Knowledge Manage Process 5:01–11.
Jensen O, Spaak E, Zumer JM (2014) Human brain oscillations: From physiological mechanisms to analysis and cognition. In: Supek S, Aine CJ, editors. Magnetoencephalography: From Signals to Dynamic Cortical Networks. Berlin Heidelberg, Berlin, Heidelberg: Springer. p. 359–403.
Jeong J (2004) EEG dynamics in patients with Alzheimer's disease. Clin Neurophysiol 115(7):1490–1505.
Jiao B, Li R, Zhou H, Qing K, Liu H, Pan H, Lei Y, Fu W, Wang X, Xiao X, Liu X, Yang Q, Liao X, Zhou Y, Fang L, Dong Y, Yang Y, Jiang H, Huang S, Shen L (2023) Neural biomarker diagnosis and prediction to mild cognitive impairment and alzheimer's disease using eeg technology. Alzheimer's Res Therapy 15(1):32.
Joachims T (1999) Making large-scale SVM learning practical. In: Schölkopf B, Burges C, Smola A, editors. Advances in Kernel Methods - Support Vector Learning. Cambridge, MA: MIT Press. p. 169–184. Ch. 11.
Jurcak V, Tsuzuki D, Dan I (2007) 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. NeuroImage 34(4):1600–1611.
Klepl D, He F, Wu M, Blackburn DJ, Sarrigiannis P (2022) Eeg-based graph neural network classification of alzheimer's disease: An empirical evaluation of functional connectivity methods. IEEE Trans Neural Syst Rehabil Eng 30:2651–2660.
Laub J, Roth V, Buhmann JM, Müller K-R (2006) On the information and representation of non-euclidean pairwise data. Pattern Recogn 39(10):1815–1826.
Lawrence N (2005) Probabilistic non-linear principal component analysis with gaussian process latent variable models. J Mach Learn Res 6(60):1783–1816.
Lawrence ND (2003) Gaussian process latent variable models for visualisation of high dimensional data. In: Proceedings of the 16th International Conference on Neural Information Processing Systems. NIPS'03. Cambridge, MA, USA: MIT Press. p. 329–336.
Lawrence ND, Quiñonero Candela J (2006) Local distance preservation in the GPLVM through back constraints. In: Proceedings of the 23rd International Conference on Machine Learning. ICML '06. New York, NY, USA: Association for Computing Machinery. p. 513–520.
Li R, Nguyen T, Potter T, Zhang Y (2019) Dynamic cortical connectivity alterations associated with alzheimer's disease: An eeg and fnirs integration study. NeuroImage: Clin 21:101622.
Lotte F, Bougrain L, Cichocki A, Clerc M, Congedo M, Rakotomamonjy A, Yger F (2018) A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. J Neural Eng 15(3):031005.
Luck S (2014) An Introduction to the Event-Related Potential Technique. A Bradford Book: Bradford Books.
Mann HB, Whitney DR (1947) On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Ann Math Stat 18(1):50–60.
Marimpis AD, Dimitriadis SI, Goebel R (2021) Dyconnmap: Dynamic connectome mapping—a neuroimaging python module. Hum Brain Mapp 42(15):4909–4939.
Matthews AGdG, van der Wilk M, Nickson T, Fujii K, Boukouvalas A, León-Villagrá P, Ghahramani Z, Hensman J (2017) GPflow: A Gaussian process library using TensorFlow. J Mach Learn Res 18 (40):1–6.
Maturana-Candelas A, Gómez C, Poza J, Pinto N, Hornero R (2019) Eeg characterization of the alzheimer's disease continuum by means of multiscale entropies. Entropy 21(6).
Micchelli CA, Xu Y, Zhang H (2006) Universal kernels. J Mach Learn Res 7:2651–2667.
Moctezuma LA, Molinas M (2020) EEG channel-selection method for epileptic-seizure classification based on multi-objective optimization. Front Neurosci 14.
Mohanty R, Sethares WA, Nair VA, Prabhakaran V (2020) Rethinking measures of functional connectivity via feature extraction. Sci Rep 10(1):1298.
Neufang S, Akhrif A, Riedl V, Förstl H, Kurz A, Zimmer C, Sorg C, Wohlschlaeger A (2011) Disconnection of frontal and parietal areas contributes to impaired attention in very early Alzheimer's disease. J Alzheimer's Dis: JAD 25:309–321.
Nunez PL, Srinivasan R (2006) Electric Fields of the Brain: The neurophysics of EEG. 2nd Edition. Oxford University Press.
Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine learning in Python. J Mach Learn Res 12:2825–2830.
Puoliväli T, Palva S, Palva JM (2020) Influence of multiple hypothesis testing on reproducibility in neuroimaging research: A simulation study and python-based software. J Neurosci Methods 337:108654.
Rasmussen CE, Williams CKI (2006) Gaussian Processes for Machine Learning. The MIT Press. p. 79–104. Ch. 4, URL http://gaussianprocess.org/gpml/.
Rastgoufard R, Alsamman A (2016) GPLVM and lava floor distance for label-deficient semi-supervised learning: Case study. In: 2016 19th International Conference on Information Fusion (FUSION). p. 84–90.
Rodriguez-Bermudez G, Garcia-Laencina PJ (2015) Analysis of EEG signals using nonlinear dynamics and chaos: a review. Appl Mathe Informat Sci 9(5):2309.
Rojas GM, Alvarez C, Montoya CE, de la Iglesia-Vayá M, Cisternas JE, Gálvez M (2018) Study of resting-state functional connectivity networks using EEG electrodes position as seed. Front Neurosci 12.
Sadaghiani S, Brookes MJ, Baillet S (2022) Connectomics of human electrophysiology. NeuroImage 247:118788.
Schölkopf B (2000) The kernel trick for distances. In: Proceedings of the 13th International Conference on Neural Information Processing Systems. NIPS'00. Cambridge, MA, USA: MIT Press. p. 283–289.
Schulz E, Speekenbrink M, Krause A (2018) A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions. J Math Psychol 85:1–16.
Shan G (2022) Monte carlo cross-validation for a study with binary outcome and limited sample size. BMC Med Inform Decis Mak 22 (1):270.
Shao J (1993) Linear model selection by cross-validation. J Am Stat Assoc 88(422):486–494.
Shirkhorshidi AS, Aghabozorgi S, Wah TY (2015) A comparison study on similarity and dissimilarity measures in clustering continuous data. PLOS One 10(12):1–20.

Stam C (2005) Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. Clin Neurophysiol 116 (10):2266–2301.
Tenenbaum JB, de Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science 290 (5500):2319–2323.
Tipping ME, Bishop CM (1999) Probabilistic principal component analysis. J Roy Stat Soc: Series B (Stat Methodol) 61 (3):611–622.
Trambaiolli LR, Lorena AC, Fraga FJ, Kanda PA, Anghinah R, Nitrini R (2011) Improving Alzheimer's disease diagnosis with machine learning techniques. Clin EEG Neurosci 42(3):160–165.
Tsoli A, Jenkins OC (2008) Neighborhood denoising for learning high-dimensional grasping manifolds. In: 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. p. 3680–3685.
Tylová L, Kukal J, Hubata-Vacek V, Vyšata O (2018) Unbiased estimation of permutation entropy in EEG analysis for Alzheimer's disease classification. Biomed Signal Process Control 39:424–430.
Tzimourta K, Afrantou T, Ioannidis P, Astrakas L, Giannakeas N (2019) Analysis of electroencephalographic signals complexity regarding Alzheimer's disease. Comput Electr Eng 76:198–212.
Urtasun R, Fleet DJ, Geiger A, Popović J, Darrell TJ, Lawrence ND (2008) Topologically-constrained latent variable models. In: Proceedings of the 25th International Conference on Machine Learning. ICML '08. New York, NY, USA: Association for Computing Machinery. p. 1080–1087.
Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, SciPy 1.0 Contributors (2020) SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 17:261–272.
Xu Q-S, Liang Y-Z (2001) Monte carlo cross validation. Chemomet Intell Lab Syst 56(1):1–11.
Xu Q-S, Liang Y-Z, Du Y-P (2004) Monte carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. J Chemom 18(2):112–120.
Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106(3):1125–1165.
Zerzucha P, Walczak B (2012) Concept of (dis)similarity in data analysis. TrAC Trends Anal Chem 38:116–128.
Zhang C, Liu C, Zhang X, Almpanidis G (2017) An up-to-date comparison of state-of-the-art classification algorithms. Expert Syst Appl 82:128–150.
Zhang H-Y, Wang S-J, Liu B, Ma Z-L, Yang M, Zhang Z-J, Teng G-J (2010) Resting brain connectivity: changes during the progress of Alzheimer disease. Radiology 256(2):598–606.
Zhao Y, Billings SA, Wei H, He F, Sarrigiannis PG (2013) A new NARX-based granger linear and nonlinear casual influence detection method with applications to EEG data. J Neurosci Methods 212(1):79–86.
Zweig MH, Campbell G (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 39(4):561–577.

APPENDIX A. EXAMPLES: EVALUATING THE APPROXIMATE SIMILARITY BETWEEN A BIPOLAR CHANNEL AND THE FUNCTIONAL NETWORKS



(Received 23 December 2022, Accepted 29 May 2023)
(Available online 8 June 2023)