

Dementia classification using a graph neural network on imaging of effective brain connectivity

Jun Cao^{a,b}, Lichao Yang^a, Ptolemaios Georgios Sarrigiannis^c, Daniel Blackburn^d, Yifan Zhao^{a,*}

^a School of Aerospace, Transport and Manufacturing, Cranfield University, Bedfordshire, MK43 0AL, UK

^b School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, UK

^c Royal Devon and Exeter NHS Foundation Trust, Exeter, EX2 5DW, UK

^d Department of Neurosciences, Sheffield Teaching Hospitals, NHS Foundation Trust, Royal Hallamshire Hospital, Sheffield, UK

ARTICLE INFO

Keywords:

EEG
Alzheimer's disease
Parkinson's disease
Directed structure learning graph neural network
Effective brain connectivity

ABSTRACT

Alzheimer's disease (AD) and Parkinson's disease (PD) are two of the most common forms of neurodegenerative diseases. The literature suggests that effective brain connectivity (EBC) has the potential to track differences between AD, PD and healthy controls (HC). However, how to effectively use EBC estimations for the research of disease diagnosis remains an open problem. To deal with complex brain networks, graph neural network (GNN) has been increasingly popular in very recent years and the effectiveness of combining EBC and GNN techniques has been unexplored in the field of dementia diagnosis. In this study, a novel directed structure learning GNN (DSL-GNN) was developed and performed on the imaging of EBC estimations and power spectrum density (PSD) features. In comparison to the previous studies on GNN, our proposed approach enhanced the functionality for processing directional information, which builds the basis for more efficiently performing GNN on EBC. Another contribution of this study is the creation of a new framework for applying univariate and multivariate features simultaneously in a classification task. The proposed framework and DSL-GNN are validated in four discrimination tasks and our approach exhibited the best performance, against the existing methods, with the highest accuracy of 94.0% (AD vs. HC), 94.2% (PD vs. HC), 97.4% (AD vs. PD) and 93.0% (AD vs. PD vs. HC). In a word, this research provides a robust analytical framework to deal with complex brain networks containing causal directional information and implies promising potential in the diagnosis of two of the most common neurodegenerative conditions.

1. Introduction

Alzheimer's disease (AD) and Parkinson's disease (PD) are two of the most common forms of neurodegenerative diseases, resulting in the loss of memory and cognitive impairments for the former while for the latter motor symptoms dominate the clinical symptomatology prior to the development of significant cognitive deficits, typically years later [1–5]. They most commonly occur among elderly people over the age of 65, but they may also affect younger individuals [6,7]. According to the World Health Organisation, more than 55 million people are currently diagnosed with a certain type of dementia. In this case, it has escalated the demand to develop efficient diagnosis methods for AD and PD.

Electroencephalography (EEG) is a non-invasive diagnostic method for studying the bioelectrical function and degeneration of the brain [8,9]. Compared with other clinical brain imaging techniques, EEG offers

higher temporal resolution at a relatively low cost [10]. The features extracted from EEG can be categorised into two groups: univariate features that are estimated from a single channel, such as amplitude and power spectrum density (PSD); and multivariate features that estimate the relationship between different regions, namely brain connectivity, which has attracted growing attention in recent years. EEG is well-accepted to estimate brain connectivity, which can reveal complex interconnections among contiguous and remote cortical regions [11–14]. Brain connectivity is believed to help people develop a deeper understanding of various neurological conditions that affect the brain and to untangle and reveal their underlying mechanisms [15–21]. It can be subdivided into neuroanatomical (or structural), functional and effective connectivity [15]. Functional brain connectivity represents the statistical dependencies between regional time series. Effective brain connectivity (EBC) tends to estimate directed causal connections among

* Corresponding author.

E-mail address: yifan.zhao@cranfield.ac.uk (Y. Zhao).

<https://doi.org/10.1016/j.combiomed.2023.107701>

Received 3 August 2023; Received in revised form 16 October 2023; Accepted 7 November 2023

Available online 15 November 2023

0010-4825/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

neural regions [22–26]. Besides the strength of the connection, the directionality of cortical information flow is considered as an additional biomarker [27,28]. The hypothesis of this study is that imaging of EBC could track more differences between AD, PD and healthy controls (HC).

One of the most popular methods for EBC identification and estimation is Granger causality (GC). GC was applied in respect to linear vector autoregressive (VAR) models of stochastic time-series data in the field of econometrics [29]. Some modifications were introduced to the original formula by Geweke [30]. Geweke's measures of Granger causality (Geweke-Granger Causality, GGC) were expanded to unconditional causality to include conditional time series components to make it possible to distinguish between direct and indirect influences [30–33]. Baccalá and Sameshima [34] introduced a novel EBC estimation method based on Granger causality - Partial Directed Coherence (PDC). It allows the estimation of causality with the PDC method using a Fourier transform of multivariate autoregressive (MVAR) model coefficients [35,36]. In other words, it operates in the frequency domain, however, it is not directly related to the power spectrum. Baccala et al. [37] introduced a modification of that method, called Generalized Partial Directed Coherence (GPDC), to improve PDC's performance in cases of noise that would cause errors in predictive modelling. Another method based on Granger causality called directed transfer function (DTF) was first introduced by Kaminski and Blinowska [38]. It applies GC on MVAR that simultaneously models all signals. Such a model, describing expressively the dataset, is transformed into the frequency domain, which can be changed into the transfer function of the system. Elements of the system indicate the causal influence from input to output in concrete frequency [39,40]. Korzeniewska et al. [41] noticed that the DTF values may be high in the parts of the spectrum where the relative power of the signal is very small, making a comparison of outflows at different frequencies. This observation led to the introduction of the full frequency Direct Transfer Function (ffDTF). Even though various EBC methods have been developed, there are limited studies to efficiently use the brain networks with directional information flows in clinical applications.

How to effectively use EBC estimations for the research of disease diagnosis is still an open problem. There are two popular strategies, including hand-crafted and imaging features [15]. On one hand, Akbarian and Erfanian [42] employed one value to represent each pair of channels by calculating DTF, PDC, and GPDC, and then performed a binary softmax classifier on the information of EBC networks. Besides, Bae et al. [43] estimated features of network properties based on Granger causality and successfully designed a support vector machine using the selected features to distinguish between alcoholics and normal subjects. In addition, a study was conducted on EBC for discrimination between Alzheimer's patients and healthy individuals using non-normalised DTF [44]. Although ffDTF was tested in this study, this version of the DTF method provides less accurate results in the classification task, compared with the original DTF. On the other hand, to obtain imaging features, Saeedi et al. [45] reconstructed a connectivity image using two connectivity measures (PDC and DTF) and eight frequency bands as the input of deep learning networks. The proposed framework achieved an accuracy of 99.24% in the case of major depressive disorder diagnosis. Besides, Olejarczyk and Jernajczyk [46] demonstrated the direction of information flow by imaging the strongest connections determined by DTF. It is found that, in patients with schizophrenia, there is a greater number of flows from the occipital to the frontal part of the brain, compared with healthy people. EBC can also play a significant role in the research of PD. Compared to healthy controls, PD patients showed reduced Theta- and Alpha-band EBC among EEG channels, which is estimated by PDC [47]. Those findings suggest that EBC has the potential to reveal abnormal cortical brain activity in patients with various neurological disorders.

To deal with complex brain networks, graph neural network (GNN) has been increasingly popular in recent years [48]. A number of studies have demonstrated that GNN outperforms other machine learning classifiers in terms of classification accuracy such as convolution neural

networks (CNN) [12,49–54]. GNN was successfully used to classify AD from health control and mild cognitive impairment [48] based on the networks estimated by graph theory. Besides, Zhong et al. [55] obtained satisfying results by combining GNN and EEG, which considers each EEG channel as a node in the graph. Brain functional connectivity was also employed to work with GNN and it obtained satisfying results in AD classification [56]. However, the existing research has not employed EBC estimations as the input of GNN. That is to say, the causal relationship has not been considered to represent the difference in dementia-related diseases. Therefore, the effectiveness of the combination of the mentioned EBC and GNN techniques for dementia diagnosis requires further investigation. The choice to use GNN in our study stems from the specific nature of our problem and the unique advantages offered by GNN over traditional methods. Our problem involves analysing complex directional relationships between brain regions, which can be well represented as graphs. GNN are specifically designed to capture intricate relationships in graph-structured data, making them well-suited for tasks where understanding the interconnections between regions is essential. Unlike traditional methods, GNN can learn and exploit these complex relationships. Besides, the proposed method allows efficiently combining brain connectivity estimations, as multivariate features, with univariate biomarkers. Furthermore, in recent years, GNN has demonstrated state-of-the-art performance in various tasks. Their success in a wide array of applications showcases their efficacy in handling diverse and complex problems.

This study proposed a framework that introduces a novel directed structure learning GNN (DSL-GNN) to leverage several EBC estimations to extract discriminative biomarkers for dementia classification. The contributions of this study can be summarised into three folds:

- A novel graph-based framework is proposed to utilise directional features from brain networks, which allows efficiently combining brain connectivity estimations, as multivariate features, with univariate biomarkers thanks to the mechanism of GNN.
- To the best of the authors' knowledge, this is the first time GNN has been applied to EBC in an attempt to distinguish AD, PD and HC. In comparison to existing GNN, the proposed DSL-GNN is capable of dealing with directional information.
- Proper visualisation is used to reveal regional information and the difference between AD, PD and HC.

2. Methods and materials

2.1. Framework

In the present paper, a novel analytical brain network imaging is constructed by a graph comprised of PSD as nodes and EBC as directional edges. We compared five EBC estimations using EEG recordings: PDC, GPDC, DTF, ffDTF and GGC. Then, a directed structure learning GNN (DSL-GNN) is developed to classify AD, PD and healthy control (HC) based on the EBC networks and frequency response of each channel. The flowchart of the proposed framework is shown in Fig. 1, including the following steps: pre-processing, feature extraction (EBC and PSD), graph construction, and DSL-GNN classification.

2.2. Preprocessing

EEG recordings were filtered with a pass band of 0.1–45 Hz using finite impulse response filtering. To reduce volume conduction effects related to the common reference electrode, 23 bipolar derivations were created from different brain regions the left frontal (F7–F3, F3–FZ), right frontal (F8–F4, F4–FZ), left frontocentral, temporo-central, temporal and centroparietal (F3–C3, T3–C3, T3–T5, C3–P3), right frontocentral, temporo-central, temporal and centroparietal (F4–C4, T4–C4, T4–T6, C4–P4), midline frontocentral, centroparietal and left and right parasagittal

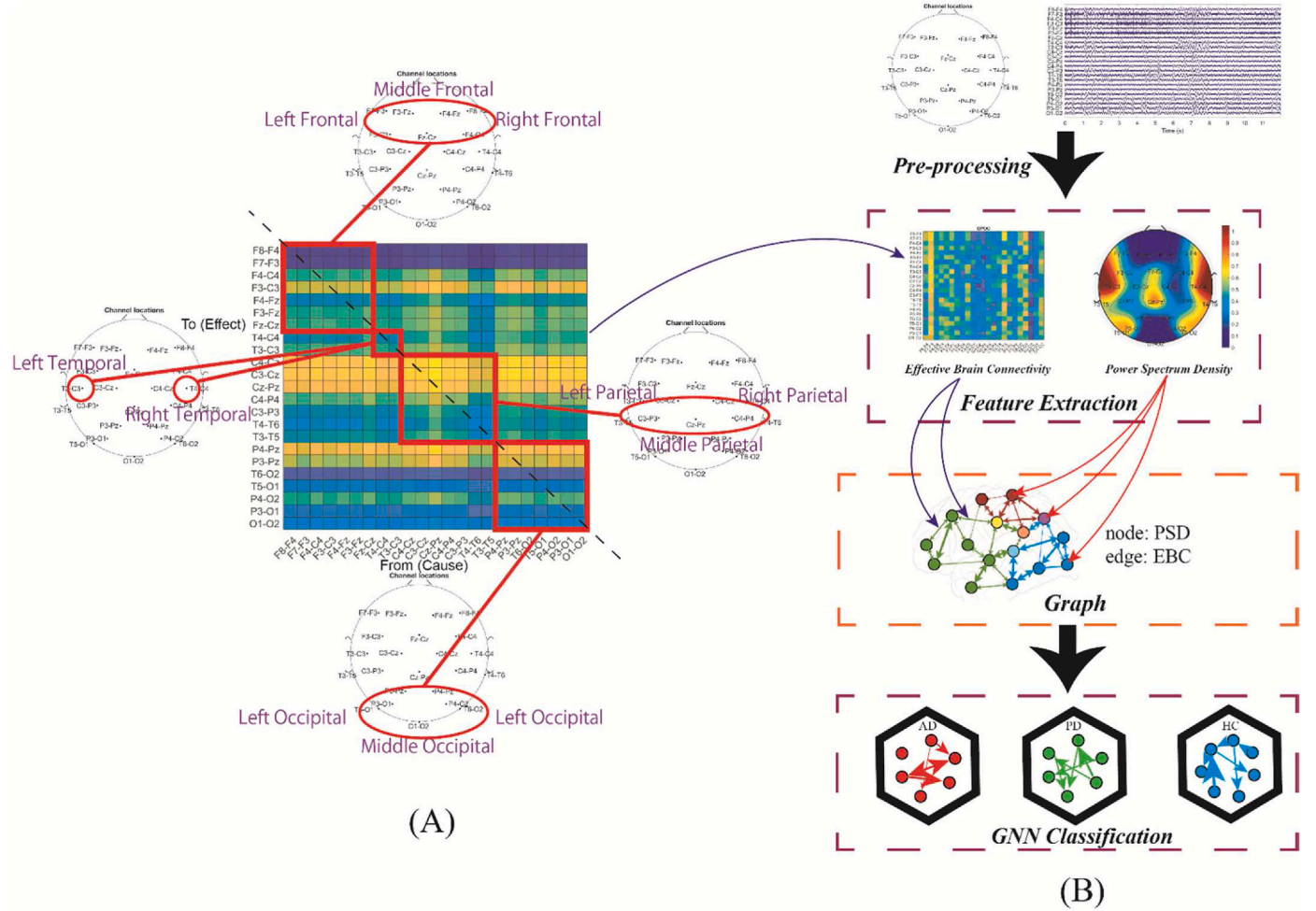


Fig. 1. (A) Effective brain connectivity visualised by a heat map marked with the corresponding brain regions; (B) the proposed framework includes four main steps: pre-processing, feature extraction (EBC and PSD), graph construction, and DSL-GNN classification. The proposed DSL-GNN employs brain networks as the input, which is constructed by a graph comprised of PSD as nodes and EBC as directional edges.

central regions (Fz-Cz, Cz-Pz, C3-Cz, C4-Cz), left parietal, parieto-occipital and temporo-occipital areas (P3-Pz, P3-O1, T5-O1), right parietal, parieto-occipital and temporo-occipital areas (P4-Pz, P4-O2, T6-O2) and midline occipital region (O1-O2)). To explore the contribution of different bands, each channel of EEG was filtered into five bands: Delta (<4 Hz), Theta (4–8 Hz), Alpha (8–12 Hz), Beta (12–32 Hz) and Gamma (>32 Hz).

2.3. Directed structure learning graph neural network

In this paper, DSL-GNN has been proposed to learn the directed connectivity from the processed connectivity map and further extract the structural feature to achieve the learning of the high-level representation. The directed graph used in this study is denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n nodes $n_i \in \mathcal{V}$, edges $(v_i, v_j) \in \mathcal{E}$. Its adjacency matrix and feature matrix are denoted as $A \in \mathbb{R}^{n \times n}$ and $X \in \mathbb{R}^{n \times c}$, respectively, where n is the number of channels which is 23 and c is the number of features of each node, which is 5 in this study. The conventional Graph Convolutional Network (GCN) [57] based methods cannot be used in the directed graph since it requests the undirected symmetric adjacency matrix. To tackle this, an approximate directed graph Laplacian method [58] based on the personalized PageRank has been introduced in this study to encode the directed features. The directed graph Laplacian L_{appr} in the symmetric normalised format is expressed as:

$$L_{appr} = I - \frac{1}{2} \left(\Pi_{appr}^{\frac{1}{2}} P_{ppr} \Pi_{appr}^{-\frac{1}{2}} + \Pi_{appr}^{-\frac{1}{2}} P_{ppr}^T \Pi_{appr}^{\frac{1}{2}} \right) \quad (1)$$

where $\Pi_{appr} = \frac{1}{\|\pi_{appr}\|_1} \text{Diag}(\pi_{appr})$. The personalized PageRank P_{ppr} is a transition matrix that includes an auxiliary node ξ to sparsify the Laplacian. It is defined as follows:

$$P_{ppr} = \begin{bmatrix} (1 - \alpha) \tilde{D}^{-1} \tilde{A} & \alpha \mathbf{1}^{n \times 1} \\ \frac{1}{n} \mathbf{1}^{1 \times n} & 0 \end{bmatrix} \quad (2)$$

where $\tilde{A} = A + I^{n \times n}$, which is the self-loops added adjacency matrix. $\tilde{D}(i, i) = \sum_j \tilde{A}(i, j)$ is the diagonal degree matrix. α is the possibility of the link between the node n and the auxiliary node ξ . π_{appr} is a part of the π_{ppr} , which is the eigenvector of the transition matrix P_{ppr} . $\pi_{ppr} = (\pi_{appr}, \pi_{\xi})$.

Based on Eq. (2), the directed graph convolution is defined as:

$$H = \text{ReLU} \left(\frac{1}{2} \left(\Pi_{appr}^{\frac{1}{2}} \tilde{P} \Pi_{appr}^{-\frac{1}{2}} + \Pi_{appr}^{-\frac{1}{2}} \tilde{P}^T \Pi_{appr}^{\frac{1}{2}} \right) X \theta \right) \quad (3)$$

where $H \in \mathbb{R}^{n \times d}$ is the output of the convolution operation with the

dimension of d . $ReLU$ is the activation function, X is the feature matrix and $\Theta \in \mathbb{R}^{c \times d}$ is the trainable weight.

Inspired by HGP-SL [59], we further learned the structure representation of the group through graph pooling that could extract the subgraph, which contains the structure information of the graph. If the feature of the node in the graph can be reconstructed by the features of its neighbourhood, such a node can be removed from the subgraph. To achieve that a criterion c is defined to calculate the score of each node for the k -th layer. It can be written as

$$c = \|L_{appr}^k H^k\|_1 \quad (4)$$

With the calculated node score, the node can be evaluated in this order. The nodes that have higher scores will be selected in the pooled subgraph. The process can be denoted as follow:

$$\begin{aligned} i &= \text{rank}(c, r \times n^k) \\ \tilde{H}^{k+1} &= H^k(i, :) \\ A^{k+1} &= A^k(i, i) \end{aligned} \quad (5)$$

where r is the pooling ratio and rank returns the indices of the required node number in the next layer, $n^{k+1} = r \times n^k$, based on the score value.

After pooling the graph into a subgraph, the new graph structure needs to be established by calculating the relationship between each pair of nodes. An attention-mechanism-based similarity score function S has been introduced to achieve this, which can be denoted as:

$$S^{k+1}(i, j) = ReLU\left(W[H^{k+1}(p, :)\|H^{k+1}(q, :)]^T\right) + \lambda A^{k+1}(p, q) \quad (6)$$

where $W \in \mathbb{R}^{1 \times 2d}$ is a trainable weight. $H^{k+1}(p, :) \in \mathbb{R}^{1 \times d}$ and $H^{k+1}(q, :) \in \mathbb{R}^{1 \times d}$ are the p -th and q -th row in the matrix H^{k+1} . $\|$ is the concatenation operation. Specifically, $A^{k+1}(p, q) = 0$, if the node n_p and n_q are not directly connected. λ is the parameter to adjust the relatively higher score of the direct connection. The similarity score function is further sparsified as S_s for the convolution operation, which can be expressed as: sparsemax is the function that could sparse the matrix with 0 and also provide non-negative and sum-to-one properties.

The structure learning convolution can be expressed as:

$$H^{k+1} = ReLU\left(\frac{1}{2}S_s^k \tilde{H}^k \Theta^k\right) \quad (8)$$

As shown in Fig. 2, the pooling operation repeats 3 times, which

provides different levels of structural information in the graph. All these levels of subgraphs need to be considered when learning the representation from the graph. Therefore, mean pooling and max pooling were used to generate a fixed size of the features. Furthermore, the features extracted from subgraphs are aggregated together for the structural representation of the graph. The generation process for the k -th layer can be expressed as:

$$g(H^k) = ReLU\left(\frac{1}{n^k} \sum_{p=1}^{n^k} H^k(p, :)\left\|\max_{q=1}^{n^k} H^k(:, q)\right\|\right) \quad (9)$$

The aggregate process can be expressed as:

$$Z = \sum_{i=1}^k g(H^i) \quad (10)$$

Finally, a classifier with MLP layers and SoftMax function is employed to achieve the classification. The prediction output Y can be expressed as:

$$Y = \text{softmax}(MLP(Z)) \quad (11)$$

2.4. Effective brain connectivity estimation

The brain networks are represented by the EBC estimation performed on the 23-channel EEG recordings, which employs five EBC techniques: PDC, GPDC, DTF, fDfDTF and GGC. The EBC of EEG tends to be visualised by a heat map. Fig. 1(A) illustrates an example of a 23*23 EBC brain network using a heat map. The colour of each kernel represents the strength of the causality from a horizontal channel to a vertical channel. The red markers indicate the regional information of the proposed visualisation with the carefully designed channel order. The principles of five employed EBC estimators are represented below.

2.4.1. Granger causality

Granger causality can estimate effective interactions from time-series data. It implements a statistical, predictive notion of causality whereby causes precede and help predict their effects [24]. It is defined in both the time and frequency domains, and it allows for the conditioning out of common causal influences. The mathematical formulation of Granger causality can be represented with a bivariate linear autoregressive model of two variables X_1 and X_2 :

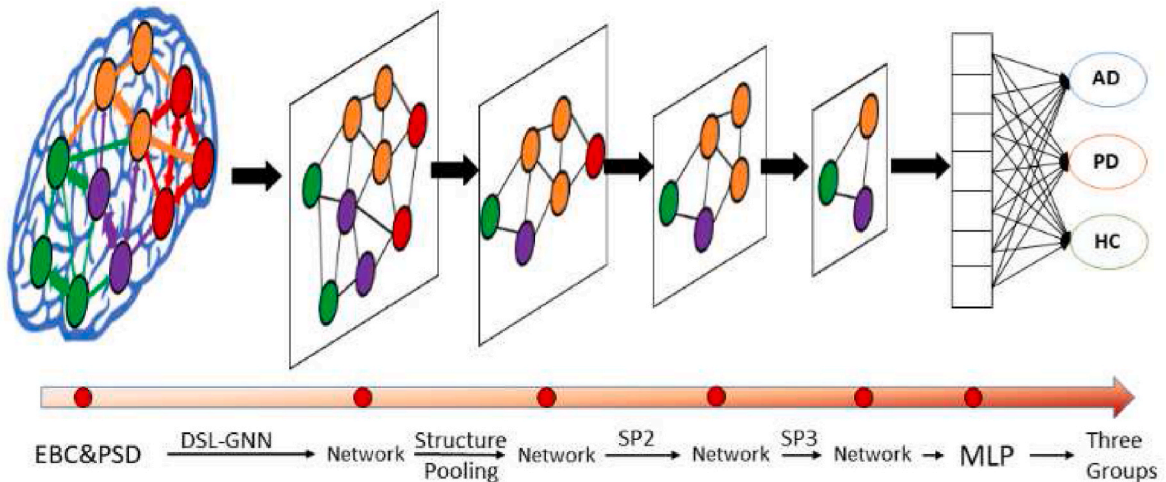


Fig. 2. The Flowchart of the proposed DSL-GNN. The input is constructed by a graph comprised of PSD as nodes and EBC as directional edges.

$$S_s^{k+1}(i, j) = \text{sparsemax}(S^{k+1}(i, j)) \quad (7)$$

$$X_1(t) = \sum_{j=1}^p A_{11,j} X_1(t-j) + \sum_{j=1}^p A_{12,j} X_2(t-j) + E_1(t) \quad (12)$$

$$X_2(t) = \sum_{j=1}^p A_{21,j} X_1(t-j) + \sum_{j=1}^p A_{22,j} X_2(t-j) + E_2(t) \quad (13)$$

where p is the maximum number of lagged observations included in the model (the model order), the matrix A contains the coefficients of the model (i.e. the contributions of each lagged observation to the predicted values of $X_1(t)$ and $X_2(t)$), and E_1 and E_2 are residuals (prediction errors) for each time series. If the variance of E_1 (or E_2) is reduced by the inclusion of the $X_1(t)$ (or $X_2(t)$) terms in the first (or second) equation, then it is said that $X_1(t)$ (or $X_2(t)$) Granger-(G)-causes X_1 (or X_2). Geweke's measures of Granger causality [30] expanded the unconditional causality to include conditional time series components to make it possible to distinguish between direct and indirect influences. Besides, GGC can be seen as Granger causality in the frequency domain [31,32,60].

2.4.2. Partial directed coherence

PDC is a method that quantifies the relation between each pair of N-channel signals, avoiding volume conduction by estimating the influences of all other signals [13]. PDC improves the concept of Partial Coherence by estimating causal influences. This method is estimated on MVAR.

$X(t)$ is a set of estimated signals from N recording channels:

$$X = [x_1(t), x_2(t), \dots, x_N(t)]^T \quad (14)$$

The MVAR process is an expressive description of the data set X :

$$\sum_{r=0}^p A(r)X(t-r) = E(t) \quad (15)$$

In this model $E(t)$ is a zero-mean multivariate uncorrelated white noise vector. $A(r)$ are the autoregressive coefficients matrix and its elements $a_{ij}(r)$ indicates the influence of $x_j(t-r)$ on $x_i(t)$ and P represents the border of the model.

PDC from the i th channel to the j th channel at the frequency f is defined as follows:

$$\varepsilon_{ij}(f) = \frac{\bar{A}_{ij}(f)}{\sqrt{\sum_{m=1}^N \bar{A}_{mj}(f) \bar{A}_{mj}^*(f)}} \quad (16)$$

where $\bar{A}_{ij}(f)$ is a frequency-domain description of $a_{ij}(r)$ [43,61]

$$\bar{A}_{ij}(f) = \begin{cases} 1 - \sum_{r=1}^p a_{ij}(r) e^{-j2\pi f r}, & \text{if } i = j \\ -\sum_{r=1}^p a_{ij}(r) e^{-j2\pi f r}, & \text{otherwise} \end{cases} \quad (17)$$

Generalized PDC (GPDC) is defined as follows by normalising (16) using standard deviation [13,62]

$$\bar{\varepsilon}_{ij}(f) = \frac{1/\sigma_i \bar{A}_{ij}(f)}{\sqrt{\sum_{m=1}^N 1/\sigma_i^2 \bar{A}_{mj}(f) \bar{A}_{mj}^*(f)}} \quad (18)$$

2.4.3. Directed transfer function

The directed transfer function (DTF) is also based on the concept of Granger causality estimated on MVAR, which models all signals simultaneously [63].

In this model $E(t)$ is a zero-mean multivariate uncorrelated white noise vector with $A(0) = 1$. $A(1)$, $A(2)$, ..., $A(p)$ are coefficients and p represents the model order. Equation (15) can be transformed into the

frequency domain which is defined as:

$$A(f)X(f) = E(f) \quad (19)$$

$$A(f) = \sum_{r=0}^p A(r) e^{-j2\pi f \Delta t r} \quad (20)$$

where.

So, $X(f)$ can be obtained by

$$X(f) = A^{-1}(f)E(f) = H(f)E(f) \quad (21)$$

$H(f)$ is the transfer function of the system and its element $H_{ij}(f)$ indicates the causal influence from the j th input to the i th output at the frequency f .

$$\beta_{ij}^2 = |H_{ij}(f)|^2 \quad (22)$$

The DTF is defined as:

Normalised DTF is defined as:

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^N |H_{im}(f)|^2} \quad (23)$$

where $\gamma_{ij}^2(f)$ describes the influence ratio of the j th channel-related cortical area on the i th channel-related cortical area with respect to the influence of all estimated cortical signals [64]. In the above formula, the denominator is dependent on frequency, so the spectral characteristics of DTF depend on the frequency properties of all EEG channels, not only the EEG input channel. To avoid this influence, the fDfDTF function was introduced. In fDfDTF, the denominator is integrated over frequencies. In this way, the spectrum of the estimator fDfDTF is not influenced by the frequency characteristics of the denominator [44].

$$\bar{\gamma}_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_f \sum_{m=1}^N |H_{im}(f)|^2} \quad (24)$$

This is an important difference between DTF and PDC since DTF is normalised for the structure that receives the signal, whereas PDC is normalised for the structure that sends the signal [43].

2.5. Dataset

Participants were recruited from Sheffield Teaching Hospitals NHS Foundation Trust Neurology clinics. Health controls (HC) were enrolled through educational meetings, word of mouth, family and friends. The Sheffield Teaching Hospital memory clinic provided the majority of these patients. This is a young-onset memory clinic seeing people predominantly aged under 65. The PD cohort was recruited during the Parkinson's UK Roadshow in York in July 2018. The remaining participants were recruited between September 2014 to December 2019. There were sixty participants in this work (70y > age > 48y) and resting state EEG recordings were undertaken from all three cohorts by our research team based at the University of Sheffield. The subjects from all groups were age-matched, HC (12 female/8 male, mean age 61y, $\pm$ SD 6.7y), AD (8 female/12 male, mean age 60y, $\pm$ SD 4.4y) and PD (10 female/10 male, mean age 60.25y, $\pm$ SD 4.96y). The EEG study underwent ethics approval by the Yorkshire and The Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). AD patients had their diagnosis confirmed between one month and up to 2 years prior to their EEG recording while they had mild to moderate cognitive deficits, according to their Mini-mental state examination. All patients and controls had brain magnetic resonance imaging (MRI) scans to eliminate other alternative causes of dementia. For the age and gender-matched HC cohort, normal MRI brain scans and cognitive assessments were required before their EEG recordings. The final diagnosis

of AD was based on the NINCDS-ADRDA criteria [65]. The diagnosis was reached based on a consensus of multidisciplinary evidence, considering clinical history, neurological examination, neuropsychological scores and neuro-radiological findings [66]. In addition, all 20 PD patients had a clinical diagnosis by a consultant with expertise in diagnosing and managing PD. All had normal cognitive function but they were in different stages of their disease. They also varied in the medication they were on and the symptoms they were experiencing.

EEG recordings were undertaken with an XLTEK 128-channel headbox (Optima Medical LTD) and Ag/AgCL electrodes at a sampling frequency of 2k Hz by implementing a modified 10-10 overlapping a 10–20 international system of electrode placement, with a referential montage (linked earlobe reference). Thirty-minute resting state (task-free – participants were instructed to rest and refrain from thinking anything specific). EEG recordings were obtained from each participant including sustained periods of keeping their eyes closed (EC) alternating with periods during which they kept their eyes open (EO). The recordings obtained were subsequently reviewed by a neurophysiologist - on an XLTEK review station - and for each participant, three 12-s artefact-free mini-epochs of EC and EO were selected. In this study, the EO and EC data are combined for analysis.

2.6. Classification

As shown in Fig. 1, the EBC and PSD features are fed into the proposed DSL-GNN to classify AD, PD and HC. To be more specific, for the Full band, the edges of DSL-GNN are represented by the EBC network (23×23) estimated by full-band EEG and the nodes are represented by the PSD values of 23 channels in five separate bands. For each band, the edges are represented by the EBC networks (23×23) estimated by corresponding-band EEG and the nodes are represented by the corresponding-band PSD values of 23 channels. There are 19 AD participants, 20 PD participants and 20 HC participants, each of which contains 6 epochs of EEG recording. Therefore, for the AD vs. HC task, the total number of data samples is $114 + 120 = 234$; for the AD vs. PD task, the total number of data samples is $114 + 120 = 234$; for HC vs. PD task, the total number of data samples is $120 + 120 = 240$; and for AD vs. PD vs. HC task, the total number of data samples is $114 + 120 + 120 = 354$. The five-fold cross-validation is employed to avoid artificial bias resulting from selecting data and more efficiently utilise the limited data [67]. In this case, 80% of the dataset is used as the training set and the remaining 20% is used as the testing set. The model training and evaluation were conducted in a workstation with the

operation system of Ubuntu 18.04. The CPU and GPU of this workstation are Intel i9-10900k and NVIDIA Quadro RTX 8000, respectively. The codes were developed in Python and the employed deep learning platform is Pytorch 1.13. For the training process, Adam was employed as the optimiser with a mini-batch size of 128. The initial learning rate and the weight decay were set as 0.001, and the number of epochs in training was set as 250.

For comparison, a Support Vector Machine (SVM) classifier is employed. To fairly reveal the performance difference, the same features extracted by EBC and PSD are fed into the proposed framework and SVM classifier, namely, 23×23 EBC features + 23×5 PSD features = 644 features for Full band; and 23×23 EBC features + 23×1 PSD features = 552 features for each of five separate bands.

3. Results

3.1. Classification performance

In this study, the proposed framework was evaluated in four classification tasks, namely, AD vs. HC, PD vs. HC, PD vs. AD and HC vs. AD vs. PD. The resulting accuracy achieved by SVM and DSL-GNN in the Full band is shown in Fig. 3 and Fig. 4 respectively using five EBC estimations individually. Overall, there is a significant improvement in accuracy (at least a 30% increase) for all four tasks using the proposed method, which suggests that DSL-GNN can simultaneously utilise both EBC (multivariate) and PSD (univariate) features more efficiently than SVM. The greatest increase occurred in the three-group discrimination, where DSL-GNN achieved 93.0% while SVM only provided 35% using both fDfTF and PSD features. As observed in Error! Reference source not found., the average performance of SVM in the three-group classification task is close to 33% while that of all other three binary classifications is slightly higher than 50%, suggesting the relatively poor performance of this approach. DSL-GNN is only limited by the worst accuracy among the binary classifications. For example, the results of the three-group classification are extremely close to the PD vs. HC task, because the accuracies of this task are lowest among the three binary classification tasks. Accuracies of DSL-GNN (96.6%) and SVM (80.0%) for classifying PD vs. AD are highest within the four tasks, which indicates that the difference between PD and AD is the most significant. It should be noted that the EBC estimation with the best performance changes from DTF for SVM to fDfTF for DSL-GNN. Furthermore, fDfTF estimation has obvious superiority over the other four estimations when DSL-GNN is employed, by producing a classification accuracy of over 93% for all four tasks.

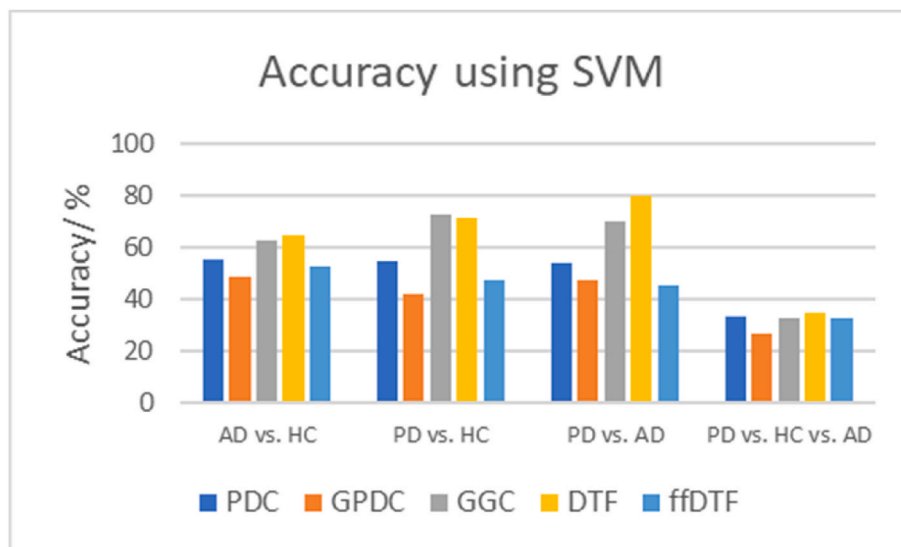


Fig. 3. Classification accuracy in Full band using SVM, employing five EBC estimations: PDC, GPDC, GGC, DTF and fDfTF.

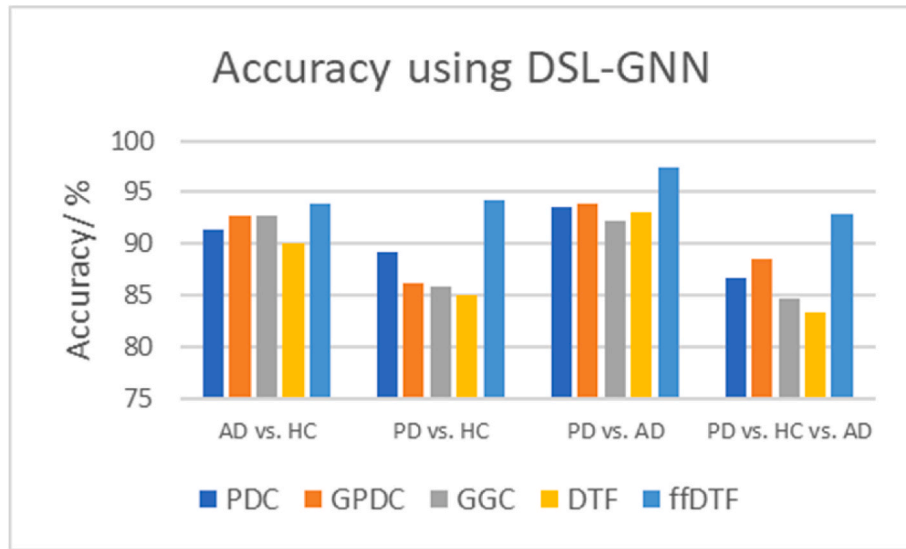


Fig. 4. Classification accuracy in Full band using DSL-GNN, employing five EBC estimations: PDC, GPDC, GGC, DTF and fDfTF.

3.2. Band contribution

To further explore the contribution of each band, the EBC and PSD features from every single band are used for comparison. Table 1 and Table 2 represent the classification accuracy based on different EBC techniques together with PSD in five separate bands, using SVM and DSL-GNN respectively. It can be found that DSL-GNN using single-band features performs much worse than the Full band due to a significant reduction of feature number, while it has smaller influences on SVM. The contribution of 5 bands to the classification outcome for SVM and DSL-GNN differs a lot. Specifically, for the AD vs. HC task, the Theta band performed better than other bands using the DSL-GNN method (about 75%), while SVM provides the highest accuracy in the Gamma band (71.4%) using the DTF technique. For the PD vs. HC task, SVM performs the greatest in the Alpha band (74.58%) using DTF, while DSL-GNN provides the highest accuracy in the Gamma band (75.83%) using GPDC. For the PD vs. AD task, the best performance is provided by DTF in the Theta band (73.94%),

while DSL-GNN achieved the highest accuracy of 84.99% in the Gamma band by employing fDfTF. For the three-group classification,

Table 2

Classification accuracy using DSL-GNN in five different bands using five EBC estimations and PSD (%).

Group	EBC Band	PDC	GPDC	GGC	DTF	fDfTF
AD vs. HC	Delta	62.43	62.70	60.99	61.50	70.99
	Theta	75.23	76.43	74.37	74.37	75.60
	Alpha	60.30	69.58	59.83	61.13	69.48
	Beta	57.69	54.61	58.49	57.32	61.10
	Gamma	49.63	58.59	51.63	51.70	67.84
PD vs. HC	Delta	50.00	58.75	49.17	49.17	67.08
	Theta	63.75	66.25	61.67	63.33	62.92
	Alpha	62.50	60.83	61.67	59.58	70.42
	Beta	50.00	54.17	50.42	50.00	63.33
	Gamma	67.92	75.83	63.33	63.75	71.25
PD vs. AD	Delta	49.63	61.17	59.18	53.81	72.40
	Theta	71.29	73.86	71.69	71.22	76.90
	Alpha	61.90	78.54	64.03	69.98	81.95
	Beta	55.55	59.46	51.34	49.86	65.48
	Gamma	59.39	79.81	54.78	56.52	84.99
PD vs. HC vs. AD	Delta	36.69	46.27	32.53	33.10	54.27
	Theta	55.91	56.53	55.05	55.59	59.31
	Alpha	45.43	53.37	44.59	42.08	52.56
	Beta	35.30	44.42	36.71	33.75	45.19
	Gamma	46.22	48.90	42.02	38.09	44.89

Table 1

Classification accuracy using SVM in five different bands using five EBC estimations and PSD (%).

Group	EBC Band	PDC	GPDC	GGC	DTF	fDfTF
AD vs. HC	Delta	56.88	51.26	61.53	51.26	49.13
	Theta	54.71	48.29	64.91	68.41	54.73
	Alpha	52.56	42.75	64.52	68.82	55.56
	Beta	62.40	44.01	62.38	57.72	51.30
	Gamma	54.31	47.42	56.85	71.41	62.01
PD vs. HC	Delta	52.92	43.33	73.75	56.25	47.08
	Theta	55.00	40.83	71.25	77.08	49.17
	Alpha	57.08	43.33	70.83	84.58	46.67
	Beta	60.83	50.00	72.92	70.00	46.25
	Gamma	57.50	44.17	68.33	65.83	48.75
PD vs. AD	Delta	54.67	51.28	69.66	67.05	48.69
	Theta	57.65	46.61	71.40	76.94	51.76
	Alpha	58.96	51.30	68.44	66.61	56.42
	Beta	62.77	48.30	70.09	70.88	52.16
	Gamma	56.40	46.61	71.80	65.83	62.00
PD vs. HC vs. AD	Delta	32.20	30.23	34.46	32.20	31.07
	Theta	31.92	26.55	37.29	31.07	31.07
	Alpha	32.20	25.99	32.20	34.75	28.53
	Beta	32.77	26.84	32.49	31.92	28.53
	Gamma	30.23	29.66	33.33	39.83	32.49

both SVM and DSL-GNN obtained unsatisfactory classification performance, even though DSL-GNN performs better than SVM with about a 15% increase in accuracy. Besides, single-band features provide much worse performance in comparison to Full-band classification with a more than 30% decrease in accuracy, while the influence on binary classification is much less. It seems that the three-group classification task requires much more features than the other three tasks. This finding may imply that the features and bands that contribute to each binary classification task vary, resulting in the Full band with more information can support better performance in three-group discrimination. In the meantime, the proposed DSL-GNN's ability to deal with a great number of features is necessary, relying on the novel graph analysis of brain networks.

3.3. Feature visualisation

The heat map (Fig. 5) and topography method (Fig. 6) are used to visualise the normalised average EBC estimations and PSD values, respectively. In Fig. 5, it is obvious that the frontal areas are rarely influenced by other regions since the fDfTF value of the first two rows is

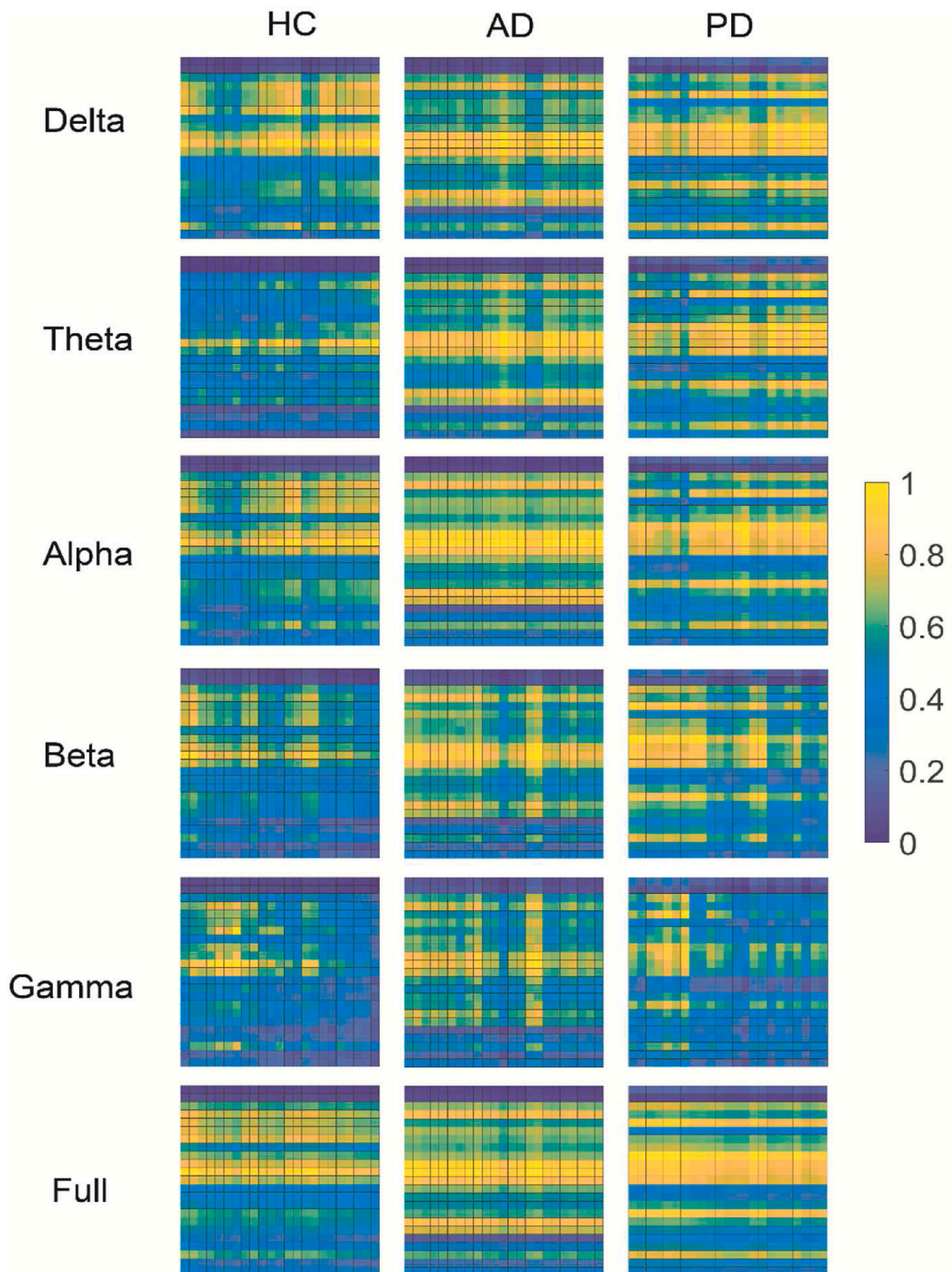


Fig. 5. The normalised average fDfTF in different band for HC, AD and PD.

extremely close to zero for three groups in every band. Compared to HC, the AD and PD groups exhibit different changes. For example, the fDfTF value of the AD group in occipital regions (Effect) increases dramatically, especially in the low-frequency bands (Delta, Theta and Alpha) from 0.5 of HC to 0.8 of AD, which suggests stronger information flow from the whole brain to the occipital area. It should be noted that the

biggest change of AD occurs in the Theta band, where the whole brain networks have a significant difference. That explains the highest accuracy of discriminating AD from HC is achieved by DSL-GNN in the Theta band to a certain extent. The left parietal and right occipital of the PD group are more causally influenced by the other regions in all bands, evidenced by the increase of fDfTF value from 0.4 to 0.8. It is worth

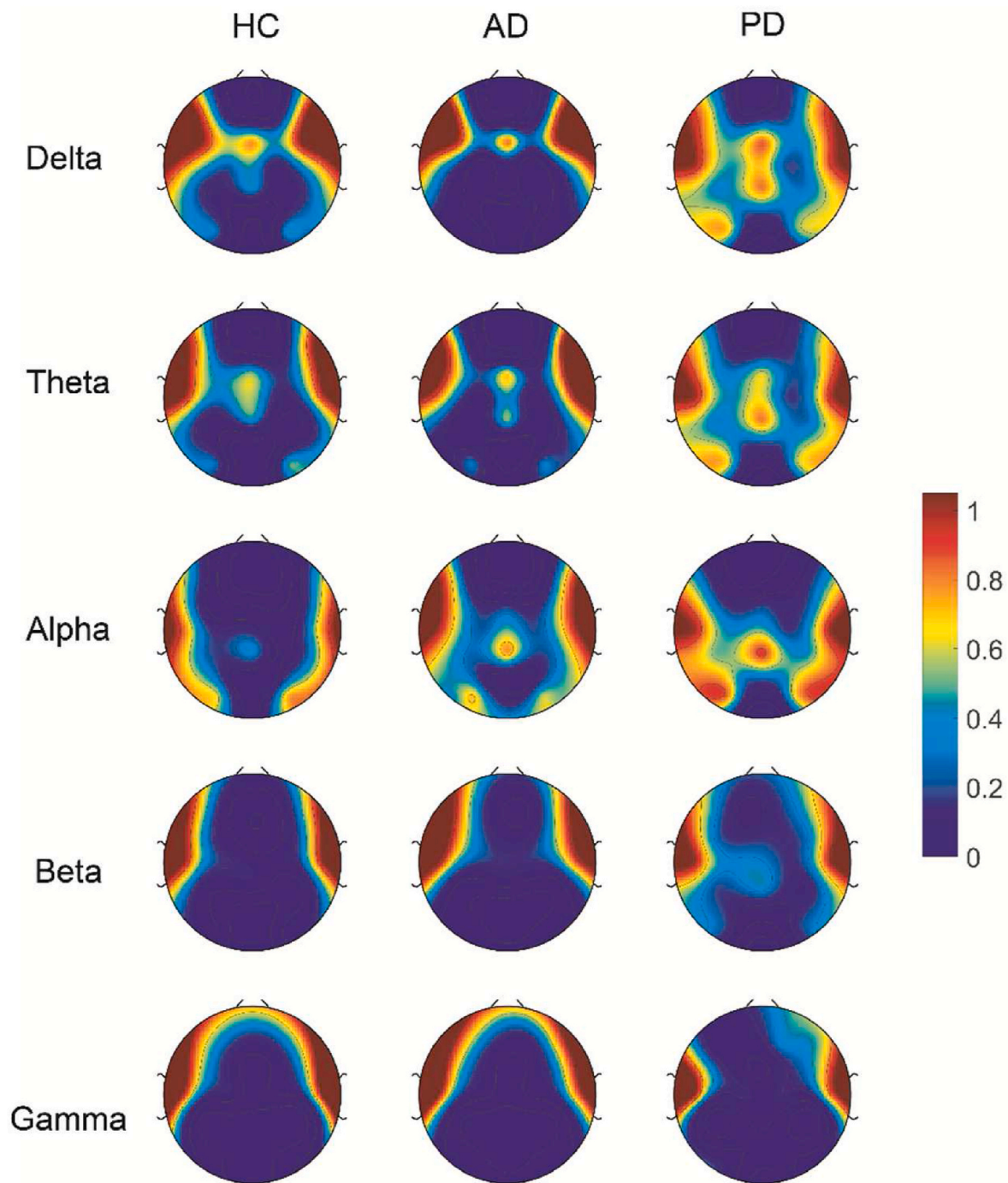


Fig. 6. The normalised average PSD in different bands for HC, AD and PD. The PSD distribution of PD shows more difference, compared with HC and AD.

noting that the bidirectional information flows within the parietal and occipital almost disappear for the PD group, which is significantly different from AD and HC. These findings agree with the classification results of PD vs. HC and PD vs. AD, which suggests the best performance is achieved in the Gamma band.

The topographic map is normalised in each band and group individually. That is to say, Fig. 6 only illustrates the different patterns of PSD distribution and it is meaningless to compare the absolute values in each topography. In Fig. 6, it is observed that the PSD distribution of PD shows more difference, compared with HC and AD. Specifically, in the low-frequency band, PD has clear higher PSD values in the parietal region of 0.8–0.9, while HC and AD range from 0.3 to 0.7. In the Gamma

band, the frontal PSD of PD (<0.1) is significantly lower than both AD and HC (>0.8). HC and AD are relatively similar in almost all bands. The greatest difference occurs in the Alpha band, in which the PSD of HC is 0.3 and AD is 0.7 in the parietal region; and PSD of HC is 0.8 and AD is 0.6 in the occipital areas. Therefore, the bands and areas that contribute the most to AD and PD classification from HC differ, and the degree of difference between AD and HC, and PD and HC varies.

4. Discussion

The novelty of this study lies in the development of a new Graph Neural Network, called DSL-GNN, to better diagnose AD by integrating

both univariate (PSD) and multivariate features (EBC) from EEG recording. The results have proven its effectiveness in improving the classification of HC, AD and PD based on the studied dataset. Table 3 compares our proposed method with the existing methods in terms of the discrimination of AD and PD from HC. The literature review suggests that very limited research focuses on the three-group classification of AD, PD and HC, although several studies obtained good results in binary classification, e.g., AD vs. HC and PD vs. HC. As cognitive dysfunctions can emerge over time in PD patients as the majority develop dementia over a decade [68] there is a clinical need in finding promising biomarkers to explore the differences between AD and PD patients. Our study not only achieved satisfying classification results in distinguishing dementia from HC, but also obtained an accuracy of 93.0% in three-group discrimination. Notably, Jeong et al. [68] performed linear discriminant analysis (LDA) on the functional connectivity estimated by Wavelet coherence and obtained approximately 80% accuracy in three binary classification tasks. However, the three-group classification was not conducted. The same features and classifier are used to deal with our dataset in the task of AD vs. PD vs. HC and the accuracy is 40.4%, which is quite similar to the results obtained by SVM using EBC in this study. The proposed framework significantly outperforms this method with an accuracy of 93.3%. Besides, Shan et al. [56] also used Wavelet coherence as the indicator, but GNN was applied to classify AD from HC only. The combination of Wavelet coherence and GNN achieved an accuracy of 92.3%, which is slightly lower than this study. Therefore, the existing research and our findings suggest that the GNN-based classifier has more capabilities in making use of numerous and complex brain network features. In the meantime, the EBC could provide more information than functional connectivity to an extent, as directionality is also used. Relying on the advancement of the proposed DSL-GNN method considering directional edges, it is not surprising that the classification performance is improved dramatically. Furthermore, to explore the effectiveness of the neural networks, Table 4 compares the recent studies that employed EEG combined with the deep learning techniques for AD or PD diagnosis. It can be found that the artificial neural network (ANN), CNN, Multilayer Perceptron Neural Network (MLP-NN) and GNN all obtained promising results in terms of classification accuracy.

The Theta band outperforms other bands in the AD vs. HC task, whichever EBC method is used. Besides, low-frequency bands (0–12 Hz including Delta, Theta and Alpha) perform better than the high-frequency band (12–45 Hz including Beta and Gamma). This finding is also supported by several previous studies [39,44,79], which suggest that the abnormalities of EEG recording at a specific frequency band (<12 Hz) are easier to be found in AD populations. Different from the AD vs. HC task, the Gamma band contributes the most to the discrimination of PD participants with relatively higher accuracy, which agrees with several recent research studies. For example, Muthuraman et al. [80]

Table 3
Performance of popular features and classifiers that are implemented on the same dataset in the discrimination task of AD or PD.

Reference	Features	Task	Classifier	Accuracy
[69]	PSD	AD vs. HC	SVM	70.2%
[70]	Mutual information	AD vs. HC	SVM	77.23%
[71,72]	Phase locking value	AD vs. HC	SVM	72.11%
[56]	Wavelet coherence	AD vs. HC	GNN	92.3%
[47]	GC	PD vs. HC	GNN	86.2%
[73]	PSD + Phase locking value	PD vs. HC	SVM	82.9%
[74]	Correlation + Coherence + Phase lag index	PD vs. HC	SVM	80.2%
[68]	Wavelet coherence	AD vs. PD vs. HC	LDA	40.4%
Our study	EBC + PSD	AD vs. HC	DSL-GNN	94.0%
		PD vs. HC		94.2%
		AD vs. PD		97.4%
		AD vs. PD vs. HC		93.0%

Table 4
Summary of the use of EEG combined with the deep learning techniques for AD or PD diagnosis in recent years.

Reference	Task	Classifier	Accuracy
[75]	AD vs. MCI vs. HC	CNN	97.1%
[76]	AD vs. HC	ANN	89.0%
[77]	AD vs. MCI vs. HC	CNN	98.9%
[78]	PD vs. HC	MLP-NN	73.5%
[2]	PD vs. HC	CNN	92.9%
Our method	AD vs. HC	DSL-GNN	94.0%
	PD vs. HC		94.2%
	AD vs. PD		97.4%
	AD vs. PD vs. HC		93.0%

found abnormal cross-frequency coupling at the Gamma band, which was not revealed within the control subject. Besides, Yuvaraj et al. [74] reported that there are no significant differences between PD patients and HC participants observed in Delta and Theta frequency bands. However, a significant difference is observed in the Beta and Gamma bands. They used four functional brain connectivity methods, including correlation, coherence, phase synchronization index, and bispectrum. Therefore, the band used to diagnose AD and PD from HC is quite different, which may explain the unsatisfying accuracy obtained by individual bands in the three-way classification task (about 50%), whereas DSL-GNN provides a highly promising accuracy of 93% using full-band features.

Although deep learning methods play an important role in the study of dementia diagnosis in recent years [81–83], effectively making use of numerous features is still challenging. Many studies suggest that more features do not always provide higher classifying accuracy [68,84]. That is to say, the best performance is obtained using a certain percentage of various features. Undeniably, the selection of features would reduce the efficiency of the classifying framework due to it is computationally demanding. Our proposed DSL-GNN method demonstrated its superiority in systematically using various brain-network markers since the mechanism of GNN supports a reasonable combination of distinct features. Specifically, the proposed framework constructs a graph using univariate features (PSD) and multivariate features (EBC) and feeds the constructed graph to the classifier. In this case, it is more robust than simply feeding all mixed features into the classifier. It learns the additional interconnection and even the direction among the nodes. That explains the GNN-based method performs significantly better than other machine learning algorithms in this study, especially in the three-group classification task in the Full band where more features are employed.

This study is noteworthy as it moves beyond a dichotomous binary perspective; namely distinguishing between the most common form of dementia (i.e., AD) and age matched healthy controls. By achieving a very high three-way classification accuracy, between HCs, AD and PD patients, our work shows the potential of our proposed EEG based framework, in detecting disease specific brain network deficits. Impressively this is highest in the distinction of participants with PD versus AD, suggesting the framework detects disease specific electrophysiological signatures. It paves the way for the development of a low cost, easy to administer, EEG-based test, for the diagnosis of the two most common neurodegenerative disorders, AD and PD, to be used in everyday clinical practice, to include primary care. This would allow early lifestyle interventions [85] and facilitate identification of suitable patients to administer newly developed disease modifying treatments for AD [86]. However, future longitudinal studies are required to determine the electroclinical associations, between the motor and cognitive deficits in the early and moderate stages of PD and AD, respectively, on one hand and the severity of EEG brain network dysfunctions, on the other. This could then translate into an objective EEG-based surrogate biomarker, of disease-specific, brain network failures, to be used in clinical trials.

In terms of generalisation and robustness of our proposed approach,

we acknowledge the validation of them is beyond the scope of the current study. We intend to explore and validate the applicability of our model in diverse applications related to neurophysiological signals and effective connectivity in future research. This article provides a promising tool (DSL-GNN) for analysing various neuroimaging modalities, such as EEG, MEG and fMRI that can be used to estimate effectivity connectivity. The satisfying results demonstrate its capability in dealing with directional networks and numerous features. In other words, our research opens up a novel direction for studies that previously lacked an effective strategy to combine various univariate and brain connectivity features. Furthermore, it would be worthwhile to investigate how our model can be utilized in various domains of EEG analysis, including but not limited to cognitive neuroscience, clinical diagnostics, brain-computer interfaces, and neurofeedback applications.

5. Conclusion

In this study, a novel DSL-GNN was developed and performed on the EBC estimations and PSD features for AD dementia classification. In comparison to the previous studies on GNN, our proposed approach enhanced the functionality of processing directional information, which builds the basis for more efficiently performing GNN on EBC. Another contribution of this study is creating a new framework for learning univariate and multivariate features simultaneously. The proposed framework and DSL-GNN are validated in four discrimination tasks and our approach exhibited the best performance with the highest accuracy of 94.0% (AD vs. HC), 94.2% (PD vs. HC), 97.4% (AD vs. PD) and 93.0% (AD vs. PD vs. HC). Additionally, the contribution of different bands is verified for these four tasks. Specifically, the most significant band varied, for example, the Theta band for AD vs. HC, the Gamma band for PD vs. HC, the Alpha and Gamma band for PD vs. AD, and the Theta band for AD vs. PD vs. HC. This research empowers the application of GNN in dementia diagnosis based on its significant outcomes. Finally, the proposed framework raises several opportunities and challenges for future research. Although the proposed DSL-GNN method has shown promising results when compared to recently published approaches, it does have certain limitations. One notable limitation is the omission of state-of-the-art features in both brain connectivity (e.g., PDC, GPDC, DTF, ffDTF and GGC) and univariate feature (e.g., PSD). Besides, this article only employs accuracy as the metric for evaluating performance, which may not provide a comprehensive understanding of the costs or consequences of associated with misclassifications. It is important to recognise that different misclassification errors can carry varying degrees of significance, and accuracy alone does not capture this important aspect. We will incorporate additional metrics such as precision, recall, and F1-score in our future studies. The proposed framework employs 5 EBC techniques as the multivariate features and PSD as the univariate features, while there are an increasing number of promising brain connectivity and single-channel biomarkers. It can be expected the performance will be further improved if more robust features are utilized.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This is independent research funded by the Royal Academy of Engineering Industrial Fellowship [#grant IF2223B-110] and carried out at the National Institute for Health and Care Research (NIHR) Sheffield Biomedical Research Centre (BRC). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

References

- [1] J.P. Amezcua-Sanchez, N. Mammone, F.C. Morabito, S. Marino, H. Adeli, A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals, *J. Neurosci. Methods* 322 (2019) 88–95, <https://doi.org/10.1016/j.jneumeth.2019.04.013>.
- [2] L. Borzi, L. Sigcha, D. Rodríguez-Martín, G. Olmo, Real-time detection of freezing of gait in Parkinson's disease using multi-head convolutional neural networks and a single inertial sensor, *Artif. Intell. Med.* 135 (2023), 102459, <https://doi.org/10.1016/j.artmed.2022.102459>.
- [3] M. Khojaste-Sarakhsi, S.S. Haghighi, S.M.T.F. Ghomi, E. Marchiori, Deep learning for Alzheimer's disease diagnosis: a survey, *Artif. Intell. Med.* 130 (2022), 102332, <https://doi.org/10.1016/j.artmed.2022.102332>.
- [4] T. Mano, K. Kinugawa, M. Ozaki, H. Kataoka, K. Sugie, Neural synchronization analysis of electroencephalography coherence in patients with Parkinson's disease-related mild cognitive impairment, *Clin. Park. Relat. Disord.* 6 (2022), 100140, <https://doi.org/10.1016/j.prdoa.2022.100140>.
- [5] W. Zhang, X. Han, S. Qiu, T. Li, C. Chu, L. Wang, J. Wang, Z. Zhang, R. Wang, M. Yang, X. Shen, Zhen Li, L. Bai, Zhuo Li, R. Zhang, Y. Wang, C. Liu, X. Zhu, Analysis of brain functional network based on EEG signals for early-stage Parkinson's disease detection, *IEEE Access* 10 (2022) 21347–21358, <https://doi.org/10.1109/ACCESS.2022.3150561>.
- [6] R. Brookmeyer, E. Johnson, K. Ziegler-Graham, H.M. Arrighi, Forecasting the global burden of Alzheimer's disease, *Alzheimer's Dement.* 3 (2007) 186–191.
- [7] R. Brookmeyer, S. Gray, C. Kawas, Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset, *Am. J. Publ. Health* 88 (1998) 1337–1342.
- [8] Y. Chen, Y. Tang, C. Wang, X. Liu, L. Zhao, Z. Wang, ADHD classification by dual subspace learning using resting-state functional connectivity, *Artif. Intell. Med.* 103 (2020), 101786, <https://doi.org/10.1016/j.artmed.2019.101786>.
- [9] G. Li, J.J. Jung, Seizure detection from multi-channel EEG using entropy-based dynamic graph embedding, *Artif. Intell. Med.* 122 (2021), 102201, <https://doi.org/10.1016/j.artmed.2021.102201>.
- [10] B. Ng, R.K. Reh, S. Mostafavi, A practical guide to applying machine learning to infant EEG data, *Dev. Cogn. Neurosci.* 54 (2022), 101096, <https://doi.org/10.1016/j.dcn.2022.101096>.
- [11] J. Cao, Y. Zhao, X. Shan, D. Blackburn, J. Wei, J.A. Erkoyuncu, L. Chen, P. G. Sarrianiannis, Ultra-high-resolution time-frequency analysis of EEG to characterise brain functional connectivity with the application in Alzheimer's disease, *J. Neural. Eng.* 19 (2022), 046034, <https://doi.org/10.1088/1741-2552/ac84ac>.
- [12] J. Cao, K. Gracjar, X. Shan, Y. Zhao, J. Zou, L. Chen, Z. Li, R. Grunewald, P. Zis, M. De Marco, Z. Unwin, D. Blackburn, P.G. Sarrianiannis, Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity, *Biomed. Signal Process Control* 67 (2021), 102554, <https://doi.org/10.1016/j.bspc.2021.102554>.
- [13] T.F. Tafreshi, M.R. Daliri, M. Ghodousi, Functional and effective connectivity based features of EEG signals for object recognition, *Cogn. Neurodyn.* 13 (2019) 555–566, <https://doi.org/10.1007/s11571-019-09556-7>.
- [14] P. van Mierlo, M. Papadopoulou, E. Carrette, P. Boon, S. Vandenberghe, K. Vonck, D. Marinazzo, Functional brain connectivity from EEG in epilepsy: seizure prediction and epileptogenic focus localization, *Prog. Neurobiol.* 121 (2014) 19–35, <https://doi.org/10.1016/j.pneurobio.2014.06.004>.
- [15] J. Cao, Y. Zhao, X. Shan, H. Wei, Y. Guo, L. Chen, J.A. Erkoyuncu, P. G. Sarrianiannis, Brain functional and effective connectivity based on electroencephalography recordings: a review, *Hum. Brain Mapp.* 43 (2022) 860–879, <https://doi.org/10.1002/hbm.25683>.
- [16] D. Goldenberg, A. Galván, The use of functional and effective connectivity techniques to understand the developing brain, *Dev. Cogn. Neurosci.* 12 (2015) 155–164, <https://doi.org/10.1016/j.dcn.2015.01.011>.
- [17] S.E. Moon, C.J. Chen, C.J. Hsieh, J.L. Wang, J.S. Lee, Emotional EEG classification using connectivity features and convolutional neural networks, *Neural Network.* 132 (2020) 96–107, <https://doi.org/10.1016/j.neunet.2020.08.009>.
- [18] W. Mumtaz, M.N.b.M. Saad, N. Kamel, S.S.A. Ali, A.S. Malik, An EEG-based functional connectivity measure for automatic detection of alcohol use disorder, *Artif. Intell. Med.* 84 (2018) 79–89, <https://doi.org/10.1016/j.artmed.2017.11.002>.
- [19] D. Wen, R. Li, M. Jiang, J. Li, Y. Liu, X. Dong, M.I. Saripan, H. Song, W. Han, Y. Zhou, Multi-dimensional conditional mutual information with application on the EEG signal analysis for spatial cognitive ability evaluation, *Neural Network.* 148 (2022) 23–36, <https://doi.org/10.1016/j.neunet.2021.12.010>.
- [20] W. Xie, R.T. Toll, C.A. Nelson, EEG functional connectivity analysis in the source space, *Dev. Cogn. Neurosci.* 56 (2022), 101119, <https://doi.org/10.1016/j.dcn.2022.101119>.
- [21] R. Zheng, Y. Feng, T. Wang, J. Cao, D. Wu, T. Jiang, F. Gao, Scalp EEG functional connection and brain network in infants with West syndrome, *Neural Network.* 153 (2022) 76–86, <https://doi.org/10.1016/j.neunet.2022.05.029>.
- [22] K.J. Friston, Functional and effective connectivity: a review, *Brain Connect.* 1 (2011) 13–36, <https://doi.org/10.1089/brain.2011.0008>.
- [23] F. He, S.A. Billings, H.L. Wei, P.G. Sarrianiannis, Y. Zhao, Spectral analysis for nonstationary and nonlinear systems: a discrete-time-model-based approach, *IEEE Trans. Biomed. Eng.* 60 (2013) 2233–2241, <https://doi.org/10.1109/TBME.2013.2252347>.
- [24] A.K. Seth, A.B. Barrett, L. Barnett, Granger causality analysis in neuroscience and neuroimaging, *J. Neurosci.* 35 (2015) 3293–3297, <https://doi.org/10.1523/JNEUROSCI.4399-14.2015>.

- [25] Yuan Wang, Yao Wang, Y.W. Lui, Generalized recurrent neural network accommodating dynamic causal modeling for functional MRI analysis, *Neuroimage* 178 (2018) 385–402, <https://doi.org/10.1016/j.neuroimage.2018.05.042>.
- [26] T.S. Zarghami, K.J. Friston, Dynamic effective connectivity, *Neuroimage* 207 (2020), 116453, <https://doi.org/10.1016/j.neuroimage.2019.116453>.
- [27] C.S. Huang, N.R. Pal, C.H. Chuang, C.T. Lin, Identifying changes in EEG information transfer during drowsy driving by transfer entropy, *Front. Hum. Neurosci.* 9 (2015) 1–12, <https://doi.org/10.3389/fnhum.2015.00570>.
- [28] G.G. Supp, A. Schlögl, N. Trujillo-Barreto, M.M. Müller, T. Gruber, Directed cortical information flow during human object recognition: analyzing induced EEG gamma-band responses in brain's source space, *PLoS One* 2 (2007), <https://doi.org/10.1371/journal.pone.0000684>.
- [29] C.W.J. Granger, Investigating causal relations by econometric models and cross-spectral methods, *Econom. J. Econom. Soc.* (1969) 424–438.
- [30] J.F. Geweke, Measures of conditional linear dependence and feedback between time series, *J. Am. Stat. Assoc.* 79 (1984) 907–915.
- [31] M. Dhamala, H. Liang, S.L. Bressler, M. Ding, Granger-Geweke causality: estimation and interpretation, *Neuroimage* 175 (2018) 460–463.
- [32] J. Geweke, Measurement of linear dependence and feedback between multiple time series, *J. Am. Stat. Assoc.* 77 (1982) 304–313, <https://doi.org/10.1080/01621459.1982.10477803>.
- [33] H. Varsheli, S.M.P. Firoozabadi, An EEG channel selection method for motor imagery based brain-computer interface and neurofeedback using Granger causality, *Neural Network* 133 (2021) 193–206, <https://doi.org/10.1016/j.neunet.2020.11.002>.
- [34] Luiz A. Baccalá, K. Sameshima, Partial directed coherence: a new concept in neural structure determination, *Biol. Cybern.* 84 (2001) 463–474.
- [35] Luiz A. Baccalá, K. Sameshima, Partial directed coherence: a new concept in neural structure determination, *Biol. Cybern.* 84 (2001) 463–474, <https://doi.org/10.1007/PL00007990>.
- [36] S. Haufe, V. Nikulin, G. Nolte, Identifying brain effective connectivity patterns from EEG: performance of Granger Causality, DTF, PDC and PSI on simulated data, *BMC Neurosci.* 12 (2011) 1–2, <https://doi.org/10.1186/1471-2202-12-s1-p141>.
- [37] L.A. Baccalá, K. Sameshima, D.Y. Takahashi, Generalized partial directed coherence, in: 2007 15th International Conference on Digital Signal Processing, 2007, pp. 163–166, <https://doi.org/10.1109/ICDSP.2007.4288544>.
- [38] M.J. Kaminski, K.J. Blinowska, A new method of the description of the information flow in the brain structures, *Biol. Cybern.* 65 (1991) 203–210.
- [39] C. Babiloni, R. Lizio, M. Marzano, P. Capotosto, A. Soricelli, A.I. Triggiani, S. Cordone, L. Gesualdo, C. Del Percio, Brain neural synchronization and functional coupling in Alzheimer's disease as revealed by resting state EEG rhythms, *Int. J. Psychophysiol.* 103 (2016) 88–102, <https://doi.org/10.1016/j.ijpsycho.2015.02.008>.
- [40] B. He, L. Astolfi, P.A. Valdes-Sosa, D. Marinazzo, S.O. Palva, C.G. Benar, C. M. Michel, T. Koenig, Electrophysiological brain connectivity: theory and implementation, *IEEE Trans. Biomed. Eng.* 66 (2019) 2115–2137, <https://doi.org/10.1109/TBME.2019.2913928>.
- [41] A. Korzeniewska, M. Mańczak, M. Kamiński, K.J. Blinowska, S. Kasicki, Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method, *J. Neurosci. Methods* 125 (2003) 195–207.
- [42] B. Akbarian, A. Erfanian, Biomedical Signal Processing and Control A framework for seizure detection using effective connectivity, graph theory, and multi-level modular network, *Biomed. Signal Process Control* 59 (2020), 101878, <https://doi.org/10.1016/j.bspc.2020.101878>.
- [43] Y. Bae, B.W. Yoo, J.C. Lee, H.C. Kim, Automated network analysis to measure brain effective connectivity estimated from EEG data of patients with alcoholism, *Physiol. Meas.* 38 (2017) 759–773, <https://doi.org/10.1088/1361-6579/aa6b4c>.
- [44] K.J. Blinowska, F. Rakowski, M. Kaminski, F. De Vico Fallani, C. Del Percio, R. Lizio, C. Babiloni, Functional and effective brain connectivity for discrimination between Alzheimer's patients and healthy individuals: a study on resting state EEG rhythms, *Clin. Neurophysiol.* 128 (2017) 667–680, <https://doi.org/10.1016/j.clinph.2016.10.002>.
- [45] A. Saeedi, M. Saeedi, A. Maghsoudi, A. Shalhaf, Major depressive disorder diagnosis based on effective connectivity in EEG signals: a convolutional neural network and long short-term memory approach, *Cogn. Neurodyn.* 15 (2021) 239–252.
- [46] E. Olejarczyk, W. Jernajczyk, Graph-based analysis of brain connectivity in schizophrenia, *PLoS One* 12 (2017) 1–28, <https://doi.org/10.1371/journal.pone.0188629>.
- [47] E. Maggioni, F. Arienti, S. Minella, F. Mameli, L. Borellini, M. Nigro, F. Cogiamanian, A.M. Bianchi, S. Cerutti, S. Barbieri, P. Brambilla, G. Ardolino, Effective connectivity during rest and music listening: an EEG study on Parkinson's disease, *Front. Aging Neurosci.* 13 (2021) 1–16, <https://doi.org/10.3389/fnagi.2021.657221>.
- [48] X. Song, M. Mao, X. Qian, Auto-metric graph neural network based on a meta-learning strategy for the diagnosis of alzheimer's disease, *IEEE J. Biomed. Heal. Inform.* 25 (2021) 3141–3152, <https://doi.org/10.1109/JBHI.2021.3053568>.
- [49] S. Anari, N. Tataei Sarshar, N. Mahjoori, S. Dorosti, A. Rezaie, Review of deep learning approaches for thyroid cancer diagnosis, *Math. Probl Eng.* 2022 (2022), <https://doi.org/10.1155/2022/5052435>.
- [50] G. Haseli, R. Ranjbarzadeh, M. Hajiaghahi-Keshteli, S. Jafarzadeh Ghouschi, A. Hasani, M. Deveci, W. Ding, HECON: weight assessment of the product loyalty criteria considering the customer decision's halo effect using the convolutional neural networks, *Inf. Sci.* 623 (2023) 184–205, <https://doi.org/10.1016/j.ins.2022.12.027>.
- [51] A.B. Kasgari, S. Safavi, M. Nouri, J. Hou, N.T. Sarshar, R. Ranjbarzadeh, Point-of-Interest preference model using an attention mechanism in a convolutional neural network, *Bioengineering* 10 (2023), <https://doi.org/10.3390/bioengineering10040495>.
- [52] R. Ranjbarzadeh, S. Jafarzadeh Ghouschi, N. Tataei Sarshar, E.B. Tirkolaee, S. Ali, T. Kumar, M. Bendeche, ME-CCNN: Multi-Encoded Images and a Cascade Convolutional Neural Network for Breast Tumor Segmentation and Recognition, *Artificial Intelligence Review*, Springer, Netherlands, 2023, <https://doi.org/10.1007/s10462-023-10426-2>.
- [53] R. Ranjbarzadeh, S. Jafarzadeh Ghouschi, S. Anari, S. Safavi, N. Tataei Sarshar, E. Babaee Tirkolaee, M. Bendeche, A deep learning approach for robust, multi-oriented, and curved text detection, *Cognit. Comput.* (2022), <https://doi.org/10.1007/s12559-022-10072-w>.
- [54] R. Ranjbarzadeh, A. Bagherian Kasgari, S. Jafarzadeh Ghouschi, S. Anari, M. Naseri, M. Bendeche, Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images, *Sci. Rep.* 11 (2021) 1–17, <https://doi.org/10.1038/s41598-021-90428-8>.
- [55] P. Zhong, D. Wang, C. Miao, EEG-based emotion recognition using regularized graph neural networks, *IEEE Trans. Affect. Comput.* (2020), <https://doi.org/10.1109/taffc.2020.2994159>, 1–1.
- [56] X. Shan, J. Cao, S. Huo, L. Chen, P.G. Sarriani, Y. Zhao, scp> Spatial-temporal graph convolutional network for Alzheimer classification based on brain functional connectivity imaging of electroencephalogram, *Hum. Brain Mapp.* 43 (2022) 5194–5209, <https://doi.org/10.1002/hbm.25994>.
- [57] A. Santerne, C. Moutou, M. Tsantaki, F. Bouchy, G. Hébrard, V. Adibekyan, J.-M. Almenara, L. Amard, S.C.C. Barros, I. Boisse, A.S. Bonomo, G. Bruno, B. Courcol, M. Deleuil, O. Demangeon, R.F. Díaz, T. Guillot, M. Havel, G. Montagnier, A. S. Rajpurahit, J. Rey, N.C. Santos, SOPHIE velocimetry of Kepler transit candidates, *Astron. Astrophys.* 587 (2016) A64, <https://doi.org/10.1051/0004-6361/201527329>.
- [58] Z. Tong, Y. Liang, C. Sun, X. Li, D.S. Rosenblum, A. Lim, Digraph inception convolutional networks, *Adv. Neural Inf. Process. Syst.* 2020-Decem (2020) 1–12.
- [59] Z. Zhang, J. Bu, M. Ester, J. Zhang, C. Yao, Z. Yu, C. Wang, Hierarchical Graph Pooling with Structure Learning, 2019.
- [60] H.E. Wang, C.G. Bénar, P.P. Quilichini, K.J. Friston, V.K. Jirsa, C. Bernard, A systematic framework for functional connectivity measures, *Front. Neurosci.* 8 (2014) 1–22, <https://doi.org/10.3389/fnins.2014.00405>.
- [61] C.W.J. Granger, N. Aug, Investigat. Causal Relat. Econometr. Mod. Cross-Spect. Methods 37 (2007) 424–438.
- [62] B. Schelter, M. Winterhalder, M. Eichler, M. Peifer, B. Hellwig, B. Guschlbauer, C. H. Lücking, R. Dahlhaus, J. Timmer, Testing for directed influences among neural signals using partial directed coherence, *J. Neurosci. Methods* 152 (2006) 210–219, <https://doi.org/10.1016/j.jneumeth.2005.09.001>.
- [63] M.J. Kaminski, K.J. Blinowska, A new method of the description of the information flow in the brain structures, *Biol. Cybern.* 210 (1991) 203–210.
- [64] F. Babiloni, F. Cincotti, C. Babiloni, F. Carducci, D. Mattia, L. Astolfi, A. Basilisco, P.M. Rossini, L. Ding, Y. Ni, J. Cheng, K. Christine, J. Sweeney, B. He, Estimation of the Cortical Functional Connectivity with the Multimodal Integration of High-Resolution EEG and fMRI Data by Directed Transfer Function, 2005, pp. 118–131, <https://doi.org/10.1016/j.neuroimage.2004.09.036>, 24.
- [65] B. Dubois, H.H. Feldman, C. Jacova, S.T. DeKosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, G. Jicha, K. Meguro, J. O'Brien, F. Pasquier, P. Robert, M. Rossor, S. Salloway, Y. Stern, P.J. Visser, P. Scheltens, Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria, *Lancet Neurol.* 6 (2007) 734–746, [https://doi.org/10.1016/S1474-4422\(07\)70178-3](https://doi.org/10.1016/S1474-4422(07)70178-3).
- [66] D.J. Blackburn, P.G. Sarriani, M. De Marco, Y. Zhao, A. Venneri, S. Lawrence, Z.C. Unwin, M. Blyth, J. Angel, K. Baster, I.D. Wilkinson, S.M. Bell, F. He, H.L. Wei, S.A. Billings, T.F.D. Farrow, 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 (2018) 1–19, <https://doi.org/10.3390/brainsci8070134>.
- [67] W. Jamal, S. Das, I.A. Oprea, K. Maharatra, F. Apicella, F. Sicca, Classification of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates, *J. Neural. Eng.* 11 (2014), <https://doi.org/10.1088/1741-2560/11/4/046019>.
- [68] D.H. Jeong, Y. Do Kim, I.U. Song, Y.A. Chung, J. Jeong, Wavelet energy and wavelet coherence as eeg biomarkers for the diagnosis of Parkinson's disease-related dementia and Alzheimer's disease, *Entropy* 18 (2016), <https://doi.org/10.3390/e18010008>.
- [69] F. Wang, Y.C. Tian, X. Zhang, F. Hu, Detecting disorders of consciousness in brain injuries from EEG connectivity through machine learning, *IEEE Trans. Emerg. Top. Comput. Intell.* (2020) 1–11, <https://doi.org/10.1109/TETCI.2020.3032662>.
- [70] J.T. Nimmy, D.P. Subha, R.N. Menon, Mutual information analysis on MCI-AD EEG signal during resting and task conditions, in: *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON 2019-October*, 2019, pp. 2295–2299, <https://doi.org/10.1109/TENCON.2019.8929242>.
- [71] M.M.A. Engels, C.J. Stam, W.M. Flier, Van Der, P. Scheltens, H. De Waal, E.C. W. Straaten, Van, Declining functional connectivity and changing hub locations in Alzheimer's disease: an EEG study, *BMC Neurol.* (2015) 1–8, <https://doi.org/10.1186/s12883-015-0400-7>.
- [72] A. Kabbara, W. El Falou, M. Khalil, H. Eid, M. Hassan, A scalp-EEG network-based analysis of Alzheimer's disease patients at rest, in: *Int. Conf. Adv. Biomed. Eng. ICABME 2017-October*, 2017, <https://doi.org/10.1109/ICABME.2017.8167549>, 0–3.

- [73] M. Hassan, L. Chaton, P. Benquet, A. Delval, C. Leroy, L. Plomhause, A.J. H. Moonen, A.A. Duits, A.F.G. Leentjens, V. van Kranen-Mastenbroek, L. Defebvre, P. Derambure, F. Wendling, K. Dujardin, Functional connectivity disruptions correlate with cognitive phenotypes in Parkinson's disease, *Neuroimag. Clin.* 14 (2017) 591–601, <https://doi.org/10.1016/j.nicl.2017.03.002>.
- [74] R. Yuvaraj, M. Murugappan, U.R. Acharya, H. Adeli, N.M. Ibrahim, E. Mesquita, Brain functional connectivity patterns for emotional state classification in Parkinson's disease patients without dementia, *Behav. Brain Res.* 298 (2016) 248–260, <https://doi.org/10.1016/j.bbr.2015.10.036>.
- [75] W. Xia, R. Zhang, X. Zhang, M. Usman, A novel method for diagnosing Alzheimer's disease using deep pyramid CNN based on EEG signals, *Heliyon* 9 (2023), e14858, <https://doi.org/10.1016/j.heliyon.2023.e14858>.
- [76] F. Duan, Z. Huang, Z. Sun, Y. Zhang, Q. Zhao, A. Cichocki, Z. Yang, J. Sole-Casals, Topological network analysis of early alzheimer's disease based on resting-state EEG, *IEEE Trans. Neural Syst. Rehabil. Eng.* 28 (2020) 2164–2172, <https://doi.org/10.1109/TNSRE.2020.3014951>.
- [77] C.J. Huggins, J. Escudero, M.A. Parra, B. Scally, R. Anghinah, A. Vitória Lacerda De Araújo, L.F. Basile, D. Abasolo, Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing, *J. Neural. Eng.* 18 (2021), <https://doi.org/10.1088/1741-2552/ac05d8>.
- [78] A.M.A. Handojoseno, J.M. Shine, T.N. Nguyen, Y. Tran, S.J.G. Lewis, H.T. Nguyen, Using EEG spatial correlation, cross frequency energy, and wavelet coefficients for the prediction of Freezing of Gait in Parkinson's Disease patients, *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* (2013) 4263–4266, <https://doi.org/10.1109/EMBC.2013.6610487>.
- [79] C. Lehmann, T. Koenig, V. Jelic, L. Prichep, R.E. John, L.O. Wahlund, Y. Dodge, T. Dierks, Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG), *J. Neurosci. Methods* 161 (2007) 342–350, <https://doi.org/10.1016/j.jneumeth.2006.10.023>.
- [80] M. Muthuraman, M. Bange, N. Koirala, D. Ciolac, B. Pintea, M. Glaser, G. Tinkhauser, P. Brown, G. Deuschl, S. Groppa, Cross-frequency coupling between gamma oscillations and deep brain stimulation frequency in Parkinson's disease, *Brain* 143 (2020) 3393–3407, <https://doi.org/10.1093/brain/awaa297>.
- [81] X. Bi, H. Wang, Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning, *Neural Network.* 114 (2019) 119–135.
- [82] C. Ieracitano, N. Mammone, A. Hussain, F.C. Morabito, A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia, *Neural Network.* 123 (2020) 176–190.
- [83] S.A.A. Shah, L. Zhang, A. Bais, Dynamical system based compact deep hybrid network for classification of Parkinson disease related EEG signals, *Neural Network.* 130 (2020) 75–84, <https://doi.org/10.1016/j.neunet.2020.06.018>.
- [84] I. Kakkos, G.N. Dimitrakopoulos, Yi Sun, J. Yuan, G.K. Matsopoulos, A. Bezerianos, Yu Sun, EEG fingerprints of task-independent mental workload discrimination, *IEEE J. Biomed. Heal. Inform.* 25 (2021) 3824–3833, <https://doi.org/10.1109/JBHI.2021.3085131>.
- [85] Y. Ko, S.M. Chye, Lifestyle intervention to prevent Alzheimer's disease, *Rev. Neurosci.* 31 (2020) 817–824, <https://doi.org/10.1515/revneuro-2020-0072>.
- [86] C.R.S. Belder, J.M. Schott, N.C. Fox, Preparing for disease-modifying therapies in Alzheimer's disease, *Lancet Neurol.* 22 (2023) 782–783, [https://doi.org/10.1016/S1474-4422\(23\)00274-0](https://doi.org/10.1016/S1474-4422(23)00274-0).



DSPACE

<https://dspace.org/>

Dementia classification using a graph neural network on imaging of effective brain connectivity

Cao, Jun; Yang, Lichao; Sarrigiannis, Ptolemaios Georgios; Blackburn, Daniel; Zhao, Yifan

2024-01-01

Elsevier

<https://doi.org/10.1016/j.compbimed.2023.107701>

Cao J, Yang L, Sarrigianis PG, et al., (2024) Dementia classification using a graph neural network on imaging of effective brain connectivity. Computers in Biology and Medicine, Volume 168, January 2024, Article number 107701