Biomedical Signal Processing and Control, Volume 67, May 2021, Article number 102554
DOI: 10.1016/j.bspc.2021.102554

# Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity

**Jun Cao[1], Kacper Grajcar[1], Xiaocai Shan[1,2], Yifan Zhao[1], Jiaru Zou[1,2], Liangyu Chen[3], Zhiqing Li[3], Richard Grünewald[4], Panagiotis Zis[4], Matteo De Marco[4], Zoe Unwin[4], Daniel Blackburn[4], Ptolemaios G Sarrigiannis[5]**

[1] School of Aerospace, Transport and Manufacturing, Cranfield University, Bedfordshire MK43 0AL, UK

[2] Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China

[3] Department of Neurosurgery, Shengjing Hospital of China Medical University, Shenyang, China

[4] Department of Neurosciences, Sheffield Teaching Hospitals, NHS Foundation Trust, Royal Hallamshire Hospital, Sheffield UK

[5] Royal Devon and Exeter NHS Foundation Trust, Exeter, EX2 5DW, UK

**Keywords: qEEG, classification, brain connectivity, correlation, coherence**

## ABSTRACT

Most seizures in adults with epilepsy occur rather infrequently and as a result, the interictal EEG plays a crucial role in the diagnosis and classification of epilepsy. However, empirical interpretation, of a first EEG in adult patients, has a very low sensitivity ranging between 29-55%. Useful EEG information remains buried within the signals in seizure-free EEG epochs, far beyond the observational capabilities of any specialised physician in this field. Unlike most of the existing works focusing on either seizure data or single-variate method, we introduce a multi-variate method to characterise sensor level brain functional connectivity from interictal EEG data to identify patients with generalised epilepsy. A total of 9 connectivity features based on 5 different measures in time, frequency and time-frequency domains have been tested. The solution has been validated by the K-Nearest Neighbour algorithm, classifying an epilepsy group (EG) vs healthy controls (HC) and subsequently with another cohort of patients characterised by non-epileptic attacks (NEAD), a psychogenic type of disorder. A high classification accuracy (97%) was achieved for EG vs HC while revealing significant spatio-temporal deficits in the frontocentral areas in the beta frequency band. For EG vs NEAD, the classification accuracy was only about 73%, which might be a reflection of the well-described coexistence of NEAD with epileptic attacks. Our work demonstrates that seizure-free interictal EEG data can be used to accurately classify patients with generalised epilepsy from HC and that more systematic work is required in this direction aiming to produce a clinically useful diagnostic method.

Published by Elsevier. This is the Author Accepted Manuscript issued with: Creative Commons Attribution Non-Commercial No Derivatives License (CC:BY:NC:ND 4.0).
The final published version (version of record) is available online at DOI:10.1016/j.bspc.2021.102554. Please refer to any applicable publisher terms of use.

## 1 INTRODUCTION

Epilepsy is one of the most common neurological disorders and it can affect people of various age, gender and ethnic origin. According to the World Health Organization in 2019, there were worldwide around 50 million people with the condition. Epilepsy is characterised by recurrent seizures and carries variable degrees of morbidity and mortality depending on the electroclinical characteristics of the seizures. Some types of epileptic seizures carry the potential to put patients and others at risk of harm, for example, if they occur during cooking, driving or swimming. Seizure detection and epilepsy diagnosis have therefore attracted many studies using various measurements. Electroencephalography (EEG) has become one of the most important approaches to assist the diagnosis of epilepsy. EEG is a non-invasive and painless technology for measuring brain activity that is also economical, easy to administer and widely available in most hospitals. When compared with other methods that provide information about the anatomical structure of the brain, such as MRI and fMRI, EEG offers ultra-high time resolution (Pievani *et al.*, 2011), which is critical to understand brain function. Synchronous networks form and dissipate in the range of 100-300ms which is thought to be the meaningful operational brain temporal scale (Varela *et al.*, 2001).

Empirical interpretation of the EEG is largely based on recognising abnormal frequencies in specific biological states (e.g. wakefulness versus sleep (Brodbeck *et al.*, 2012; Lioi *et al.*, 2017)), the spatial-temporal and morphological (e.g. sharp waves, spikes etc.) characteristics of paroxysmal (Dash *et al.*, 2018) or persistent discharges (Renzel *et al.*, 2017), reactivity to external stimuli and activation procedures (like a period of hyperventilation (Watanabe *et al.*, 2018) or intermittent photic stimulation (Visani *et al.*, 2010)). Despite being useful in many instances, these practical approaches for interpreting EEGs leave important linear or nonlinear interactions between various brain network anatomical constituents, buried undetected within the recordings as such interactions are far beyond the observational capabilities of any specialised physician in this field (Sarrigiannis *et al.*, 2014, 2018; Blackburn *et al.*, 2018). Another limitation is that previous studies are based on univariate methods, focusing on single EEG recording channels. Studying brain functional connectivity among multiple channels (i.e. multivariate methods), by examining the magnitudes of temporal correlations or coherence with frequency, is increasingly being recognised as an important approach with many advantages. Various abnormalities in brain functional connectivity have been reported for numerous brain disorders (Uhlhaas and Singer, 2006; Guevara Erra, Perez Velazquez and Rosenblum, 2017) but to the best of our knowledge, this has not yet translated in a diagnostic method for clinical use. More efforts are required to further explore the brain connectivity measures as new biomarkers for the diagnosis of epilepsy. Different methods were previously introduced to measure brain connectivity, based on linear or nonlinear association in the time, frequency and time-frequency domains. As the most widely studied type of brain connectivity, functional connectivity can be presented by magnitude squared coherence (Sakkalis, 2011; Battaglia and Brovelli, 2019; Tafreshi, Daliri and Ghodousi, 2019), Minimum Description Length (Sakkalis, 2011; Salman, Grover and Shankar, 2018), phase synchronisation (Sakkalis *et al.*, 2006), phase-locking value (Lotte *et al.*, 2018; Bedo, Ribary and Ward, 2020), robust synchronisation (Sakkalis *et al.*, 2006; Delgado-Restituto, Romaine and Rodríguez-Vázquez, 2019), non-linear correlation coefficient (Wendling *et al.*, 2010), correlation (Horstmann *et al.*, 2010) and power distributions (Rosch *et al.*, 2017) etc. The prospect to use those techniques to aid the diagnosis of epilepsy remains unknown.

Fortunately, the large majority of patients with epilepsy experience seizures only infrequently, although there are exceptions to this rule, for example, the type of seizures characterizing children with various forms of absence seizures; absences can frequently occur daily and can, in addition, be easily

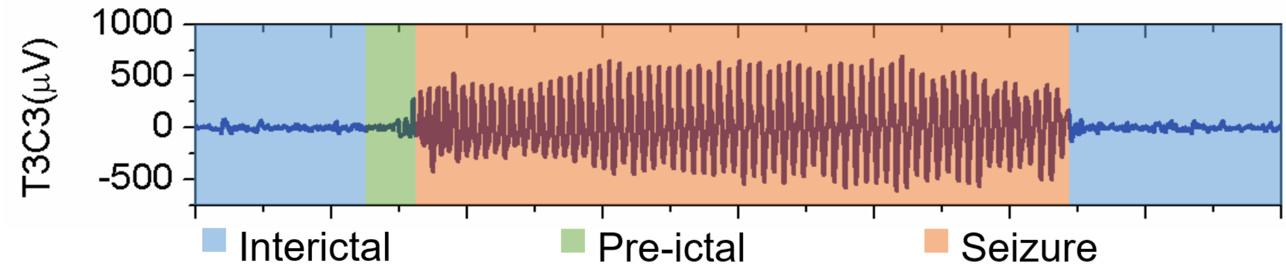



Fig. 1. A typical EEG recording before during and after a generalised epileptic seizure (typical absence) recorded from the left centrotemporal area. The ictal segment (i.e. seizure)) contains large magnitude epileptic discharges, which can be easily distinguished from the seizure-free recording (interictal).

precipitated by a brief period of hyperventilation. The EEG recording of a typical absence seizure is shown in Fig. 1 and can be separated into three states: seizure (ictal), pre-seizure (pre-ictal) and seizure-free (interictal). The difference between a seizure and interictal EEG segment is clearly visible in terms of the amplitude and morphology of the generalised epileptiform discharges dominating the ictal phase. Thus, epilepsy diagnosis when one of the patient's typical seizures is captured can be relatively straightforward. Nonetheless, as most seizures in adults with epilepsy occur rather infrequently, the interictal EEG (i.e. recorded while the patient is not experiencing any epileptic seizures) is the most commonly available data for clinicians attempting to diagnose epilepsy. The interictal epileptiform discharge (i.e. sharp wave and/or spike) recognised empirically by the reporting EEG physician, is an expression of the abnormal neuronal and brain network behaviour, a demonstration of cortical hypersynchrony and hyperexcitability. However, a first EEG in adult patients, subsequently proven to have epilepsy, has very low sensitivity, ranging between 29-55%, that can go up to 80 to 90% on repeats of the examination (Pillai and Sperling, 2006).

EEG is one of the most useful diagnostic procedures for epilepsy. It provides evidence for the diagnosis, classification and management of different types of epilepsy (Smith, 2005; Noachtar and Rémi, 2009). It should be noted that most of the current research in this topic focus on seizure detection, in subjects that have epilepsy while relatively limited work is centred on normal in appearance, interictal EEG epochs in patients where there is independent strong evidence for the electroclinical diagnosis of epilepsy. Amin *et al*. (2020) proposed a novel method based on wavelet analysis and arithmetic coding to achieve efficient classification between epileptic seizure signals and seizure-free signals, while our paper tended to classify epilepsy patients from healthy people merely using seizure-free EEGs. Similarly, to identify epileptic seizures from EEG signals, Dhiman & Priyanka (2014) proposed a novel scheme based on discrete wavelet packet transform and Zhu *et al.* (2014) proposed a fast weighted horizontal visibility graph constructing algorithm. In studies aiming to diagnose epilepsy, most papers use data that include ictal EEG recordings, such as (Fani, Azemi and Boashash, 2011; Xie and Krishnan, 2013; Kaya and Ertuğrul, 2018; Vijay Anand and Shantha Selvakumari, 2019; Akbarian and Erfanian, 2020), where seizure detection is also involved. However, analysis of interictal seizure-free EEG has gradually attracted more attention recently. For instance, Lopes *et al*. (2019) suggested that the interictal EEG supported by many publications can also contribute to epilepsy diagnosis and they developed a framework to classify focal and generalized epilepsy by extracting Ictogenic Spread as biomarker from normal EEG. Horstmann *et al*. (2010) found differences in the functional networks of seizure-free intervals of patients with focal epilepsy (treated with anti-epileptic drugs) and healthy controls as measured from EEG/MEG. Besides, the existence of alpha rhythm abnormalities in patients with epilepsy was explored based on the information obtained from interictal scalp EEG recordings (Pyrzowski *et al.*, 2015). In the appropriate clinical context, the presence of a generalised interictal epileptiform discharge (IED) further strongly supports the clinical diagnosis of generalised epilepsy which will be the focus of this work.

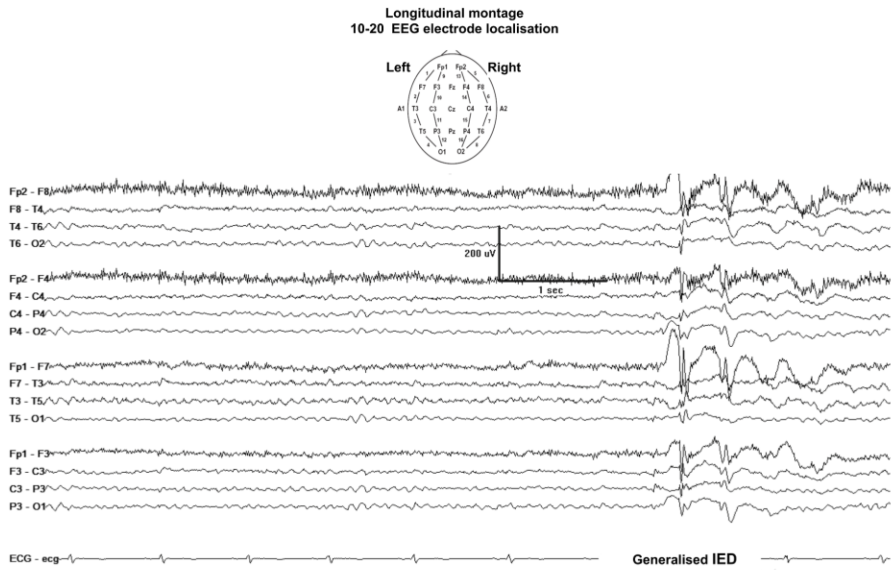



Fig. 2. EEG recording from one of the patients of the epileptic cohort showing an interictal epileptiform discharge (IED). This demonstrates a clear generalised distribution in keeping with the patient's diagnosis of generalised epilepsy. Of note, the EEG is "normal-looking",i.e.empirically classified as normal, prior to the generalised paroxysmal discharge. "Normal-looking" EEG epochs were the dominant feature on this interictal EEG recording, a common occurrence in this group of patients.

The research question of this study is to examine if appropriate analysis of "normal-looking", i.e. empirically classified as normal, interictal EEG data can identify adult patients with generalised epilepsy. This task is unlikely to be solved by direct inspection of EEG recordings due to the lack of obvious patterns to distinguish patients with epilepsy from age and gender-matched healthy controls. Achieving this aim could greatly improve the diagnostic sensitivity and specificity of interictal EEG, reducing diagnostic times and cost while ensuring that epileptic patients can promptly initiate appropriate treatment. Furthermore, the EEG recording time and repeat EEG recordings can be kept to a minimum leading to much more efficient use of recourses and increased patient satisfaction. Another challenge of this study was to determine whether the "normal-looking" interictal EEG recordings of patients with non-epileptic attacks (the ictal EEG is normal during those episodes) occurring within the context of a psychogenic disorder, called non-epileptic attack disorder (NEAD) could be differentiated from equivalent recordings of patients with epilepsy. NEAD is a brain-related disorder which involves psychogenic non-epileptic seizures (Sheldon and Agrawal, 2019). According to Milán-Tomás *et al*. (2018), about 20% to 40% of patients diagnosed with epilepsy also have NEAD rendering differentiation between the two conditions a challenging task for physicians. Furthermore, frequently NEAD patients receive several anti-epileptic drugs (AED) due to their seizures being pharmacoresistant as they are psychogenic in nature. As a result misdiagnosis of NEAD may cause patients serious iatrogenic arrhythmia (Brown *et al.*, 2011). Therefore, a computer-aided classification between NEAD and epilepsy becomes important but very challenging as the two conditions can co-exist in the same subject and appropriate data labelling is problematic.

Different from most of the existing studies either focus on single EEG channel or EEG data with seizures, this paper proposes a novel framework to diagnose epilepsy based on interictal seizure-free EEG data only using a set of estimates of linear and nonlinear brain functional connectivity. A machine learning approach is then employed to classify (a) the epilepsy group (EG) vs healthy controls (HC), and (b) EG vs NEAD, followed by a visualisation of classification results. Although the methods to calculate the connectivity are not new, the attempt to systematically evaluate their potential on interictal seizure-free EEG data in time, frequency and time-frequency domains has not been done before, particularly for the challenging classification of EG vs NEAD.

# 2 METHODOLOGY

## 2.1 Dataset

In this study we have retrospectively selected video EEGs (vEEGs) with occasional generalised IEDs, providing electrophysiological evidence for the diagnosis of epilepsy. All selected patients were isolated from the Royal Hallamshire Hospital (Sheffield, UK) Department of Neurophysiology database with the following inclusion criteria: standard interictal EEG available containing at least one well defined generalised IED (Fig. 2); age between late teens to 61 years (based on a cohort of HCs available from previous work to ensure no significant age differences between groups occurred); their EEGs included periods of wakefulness with eyes open (EO) and eyes closed (EC) epochs; previous history of at least one witnessed generalised tonic-clonic seizure without any other known type of seizures. We have also selected a cohort of patients with NEAD where we have captured at least one typical psychogenic attack on their vEEG recording for which there was no evidence of ictal or IED, no dynamically evolving ictal EEG patterns and no clear history of other types of seizure. The following exclusion criteria were also applied: learning difficulties; sleep deprivation the night before the EEG was recorded; known history of drug addiction; refractory epilepsy; any other known neurological disorder other than epilepsy or NEAD. Some of the NEAD patients received various AEDs but there was no convincing evidence on their past medical history (reviewed during their EEG recording) to suggest epileptic seizures. However admittedly this cannot be entirely excluded for this group of patients.

A Natus Headbox (Optima Medical, Ltd.) at a sampling rate of 500Hz (analogue bandwidth 0.1–200Hz) and a standard international 10–20 system of electrode placement positions were used for the recordings for all subjects. The EEG data was recorded from the standard 21 electrodes of the 10-20 system of electrode placement. The investigated EEG data comes from three groups: HCs, EG and NEAD. 10 HCs (6 females, mean age 37±15y), 15 EG (10 females, mean age 33±12y) and 14 NEAD cases (10 females, mean age 33±13y) were collected, details of which can be seen in Table S1 of Supplementary material. It should be noted that EG and NEAD participants were consecutively selected for the best possible match of age to our healthy control cohort. Additionally, the data for each participant was divided into two states: eyes open (EO) and eyes closed (EC). For each HC participant in each eye state, 3 trials were collected and each epoch lasts 12 seconds. For each NEAD and EG participant in each eye state, 2 trials were collected and each epoch lasts 12 seconds. The total number of available data for each group is therefore similar to ensure the fairness of training and validation, as shown in Table S1. All HC participants provided informed consent as part of a project approved by the Yorkshire and the Humber (Leeds West) Research Ethics Committee (reference number 14/YH/1070). For the retrospectively selected EEG data for the NEAD and EG cohorts, ethics approval for use of patients' EEGs for the development of novel qEEG methodologies was granted both from the University of Sheffield and the NHS ethics committees (SMBRER207 and 11/YH/0414).

Bipolar and unipolar (i.e. referential) derivations are the two main types of EEG recordings used in everyday clinical practice. The unipolar montage uses one channel as a source and the other as a reference that is usually fixed; this is the default mode of recording EEG data in routine clinical practice. The bipolar montage is obtained by subtraction of two unipolar derivations. This study used bipolar derivations to minimise the effects of volume conduction introduced by a common reference (Fein *et al.*, 1988; Nunez *et al.*, 1997). We have produced 23 bipolar channels calculated by the 21-channel unipolar recordings after excluding the frontopolar electrodes (Fp1 and Fp2) due to their vicinity to the eyes that results in high levels of eye blink artefacts.

The EEG epochs were collected with Spike2 (version 9) software where data filtering and labelling was also undertaken. The filtering method was performed using the below equation

$$\boldsymbol{y(k) = x(k) - \frac{1}{2c\sum_{i=k-c/f_s}^{k+c/f_s} x(i)}} \tag{1}$$

where $x$ is the input EEG signal, $y$ is the output signal, $k$ is a discrete-time point, $f_s$ is the sample rate and $c$ is a time constant value. The value of $c$ is set at 0.2s. The filtering using a time constant results in a high pass filtering of the signal where the cut-off frequency $f_c$ is:

$$\boldsymbol{f_c = \frac{1}{2\pi c}} \tag{2}$$

The value of $f_c$ is equal to 0.79Hz.

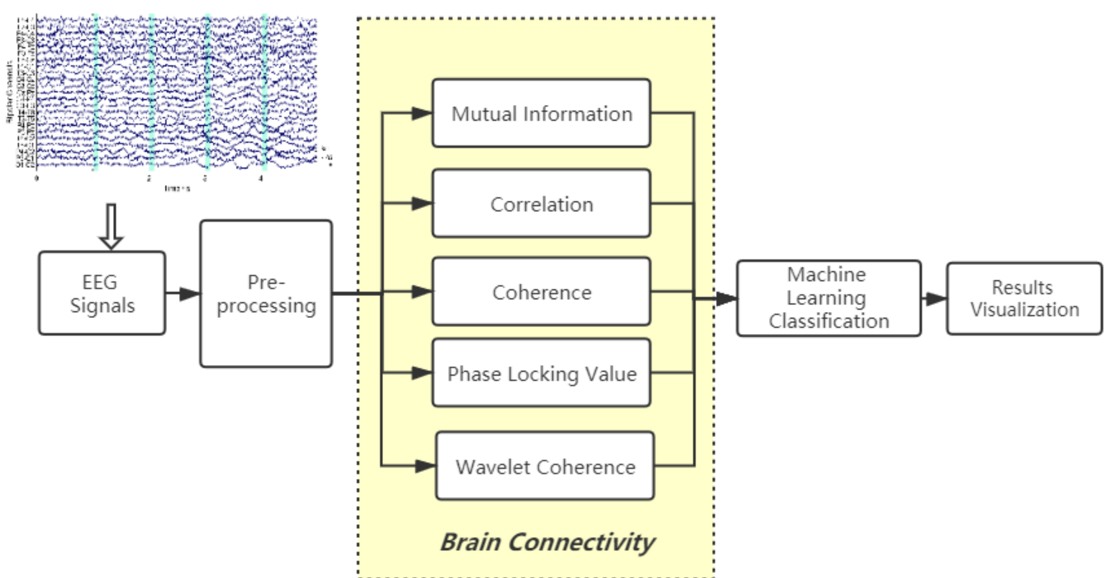



Fig. 3. Flow Diagram for EG vs. HC and EG vs. NEAD classification

All data were selected by a specialised physician in clinical neurophysiology to ensure no interictal EEG abnormalities were included. Additionally, care was taken to select relatively artefact free epochs. Furthermore, Fig. 3 shows the proposed framework, including pre-processing, features extraction, classification and result visualization, details of which are explained below. Sensor level brain functional connectivity measures, estimated by five methods, were extracted from EEG data as features for further classification.

## 2.2 Data pre-processing

Before estimating functional connectivity, data was pre-processed by the following steps:

1. For the groups of EG (15 subjects) and NEAD (14 subjects), there are 2 epochs for EO and EC for each subject. For the HC group (10 subjects), there are 3 epochs for EO and EC for each participant. To increase the number of samples for classification, each dataset (12 seconds) was segmented into 3 mini-epochs, 4 seconds each. In total, there were 528 labelled samples for 3 groups (EG:180, NEAD:168 and HC:180), prepared for feature extraction and classification. As brain interactions are usually highly dynamic, prolonged EEG segments can obscure this important characteristic (Durongbhan *et al.*, 2019).

2. Each segment of data was normalised in the range from -1 to 1 for all bipolar channels. This step was performed mainly for visualisation purposes at the beginning of this work to search for any possible artefacts and does not affect any of the connectivity measures used in this paper.

3. In this framework, it is tended to extract features in each frequency band. Therefore, each data segment was filtered to produce six frequency bands: full band (no filtering), Alpha (8-15Hz), Beta (15-32Hz), Gamma (>32Hz), Delta (<4Hz), Theta (4-8Hz). This operation was performed using an FIR filter with an order of 600. The frequency ranges of the Delta and full bands were modified to eliminate the effect of the time constant filtering.

## 2.3 Functional connectivity estimation

This paper implements 5 different measures in time, frequency and time-frequency domains to represent the brain functional connectivity between two EEG signals $x_i$ and $y_i$. These measures with their extension are then constructed as classification features. The properties of these measures are presented in Table 1.

### *2.3.1 Mutual Information*

Mutual Information (MI) measure indicates the mutual dependence of two signals, i.e. how much information is shared between two signals (Cover and Thomas, 2005). It is based on a probability function and entropy. The entropy of a signal $X$ with a length of $n$ is expressed as

$$\boldsymbol{H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i)} \tag{3}$$

where $p(x_i)$ is the probability function and values $x_i$ ($i = 1,2,3, \dots, n$) represent all possible values of the signal $X$. If the entropy of the signal is high, then the signal contains a lot of different values. For low entropy, the signal is more organised, for example, more values are repeated in the signal. A joint probability $p(x_i, y_j)$ is the probability that $X = x_i$ and $Y = y_j$. The mutual information $I(X;Y)$ between two signals $X$ and $Y$ is expressed as

$$\boldsymbol{I(X;Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i) p(y_j)}} \tag{4}$$

where n and m are the length of the signal X and Y respectively. A total of 256 bins were used to calculate the value of mutual information.

### *2.3.2 Correlation*

This well-known measure shows how two signals are correlated with each other corresponding to a time shift between these two signals. In this study, the cross-correlation used is calculated using the following formula

$$\boldsymbol{R_{xy}}[\mathbf{n}] = \frac{\sum_{k=-\infty}^{\infty} x[k]y[k+n]}{\sqrt{\sum_{k=-\infty}^{\infty} x[k]^2} * \sqrt{\sum_{k=-\infty}^{\infty} y[k]^2}} \tag{5}$$

where $x[k]$ and $y[k]$ are discrete signals and $n$ is the lag index. Cross-correlation is a measure of the linear correlation (dependence) between two signals, giving a value between -1 and 1 inclusive, where 1 indicates a total positive correlation, 0 indicates no correlation, and -1 indicates a total negative correlation. Three features are taken from the cross-correlation, namely the maximum value of correlation (CorrMax), the mean value of correlation (CorrMean) and the correlation lag at the maximum value of correlation (CorrLag). It should be noted that the absolute value of the correlation is taken into account in terms of the maximum value of the correlation. This is due to the fact that two EEG signals can also be negatively correlated.

### *2.3.3 Coherence*

The coherence in the frequency domain is estimated using magnitude squared coherence (MSC), expressed as

$$\boldsymbol{C_{xy}(f)} = \frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)} \tag{6}$$

where $G_{XY}(f)$is the cross-spectral density of signals $X$ and $Y$. $G_{XX}(f)$ and $G_{YY}(f)$ are the auto-spectral density of these signals. The step size of the frequency is $f_S/nFFT = 500/2048 = 0.2441$ Hz. Two features are taken from the MSC: maximum (CohMax) and mean values (CohMean). Although these features are extracted for six frequency bands, MSC is computed only for signal filtered to the full band range (2Hz-60Hz). To obtain MSC for different bands, MSC is segmented using the appropriate frequency ranges for each band and the mean and maximum values of the MSC are taken from these frequency segments.

### *2.3.4 Phase Locking Value*

The Phase Locking Value (PLV) was first introduced by Lachaux *et al*. (Lachaux *et al.*, 1999) in 1999. PLV measures the significance of the phase covariance between two signals. It is defined as

$$\boldsymbol{PLV}(\mathbf{t}) = \frac{1}{N}\left|\sum_{N=1}^{N} e^{i\theta(t,n)}\right| \tag{7}$$

where $\theta(t,n)$ is the difference in the phase of two signals. PLV values are in the range from 0 to 1. If the phase difference of two signals remains the same, PLV is close to one. A PLV close to zero indicates that there is no phase synchrony between two signals.

To compute the instantaneous phase of each EEG signal, a Hilbert transform is used. For the signal $x(t)$, this transform allows creating an analytic signal $a_1(t)$ as:

$$\boldsymbol{a_1(t) = x(t) + \gamma h_1(t)} \tag{8}$$

where $\gamma$ is an imaginary number and $h_1(t)$ is the Hilbert transform of $x(t)$. The instantaneous phase is defined as an angle between the real and imaginary part of the analytic signal and is written as

$$\boldsymbol{\theta_1(t) = \mathbf{arctan}\left(\frac{h_1(t)}{x(t)}\right)} \tag{9}$$

The same method can be applied to the signal $y(t)$. Lachaux *et al*. (Lachaux *et al.*, 1999) used $N$ trials to obtain statistically significant values of PLV. This stage is omitted in this project, thus $N = 1$ during the calculation of the PLV. Additionally, the differences between the instantaneous phases are computed for all points in time and the differences are summed to obtain the PLV. To normalise the PLV, the obtained sum is divided by the signal length. The final equation to calculate PLV is then rewritten as

$$\boldsymbol{PLV = \frac{1}{T}\left|\sum_{i=1}^{T} e^{i\theta(t_i)}\right|} \tag{10}$$

where $T$ represents the size of a time window and $t_i$ is a discrete point in time.

*2.3.5 Wavelet Coherence*

The wavelet formulation of coherence between two signals, $x$ and y, and in the frequency w and time t domain, can be formulated as (Zhao *et al.*, 2018):

$$coh_{xy}^2(w,t) = \frac{\left|\, S_{xy}(w,t)\right|^2}{S_x(w,t)\, S_y(w,t)}$$

$$\boldsymbol{S_{xy}(w,t) = \mathbb{E}\left(W_x(w,t)\overline{W_y(w,t)}\right)} \tag{11}$$

where $S_{xy}(w,t)$ is the wavelet cross-spectrum between $x$ and y and $S_x(w,t)$, $S_y(w,t)$ are the corresponding auto-spectrums. Working with two single signals (single realisation) usually requires using a smoothing operator (see $f(\cdot)$ operator in Eq. (12)), and ergodicity properties should be assumed (Sairamya *et al.*, 2018).

$$\boldsymbol{\widehat{coh_{xy}^2}(w,t) = \frac{\left|f\left(S_{xy}(w,t)\right)\right|^2}{f\left(S_x(w,t)\right)\cdot f\left(S_y(w,t)\right)}} \tag{12}$$

Two features are extracted from the wavelet coherence: mean value (WCohMean) and maximum value (WCohMax). As wavelet coherence is in the time-frequency domain, mean and maximum values are taken from all time points and the relevant frequency ranges. The mother wavelet *Morlet* was used for this study.

## 2.4 Classification using machine learning

Considering the estimated measures of functional connectivity between two signals, with the 23 channels (shown in Table S2 of the supplementary material), used in this study, there are 253 possible combinations ($C_2^{23}$) when any 2 bipolar derivations are paired together. These were organised in a pairwise manner by taking the first bipolar derivation in the list (F8-F4) and pairing it with every other bipolar derivation according to their order on the list (F8-F4:F7-F3, F8-F4:F4-C4, F8-F4:F3-C3, …).

The process was repeated for all other channels until the end of the list. However, since each channel is bipolar in nature, any pair with common electrode locations (such as F8-F4 and F4-C4) is neglected as this could lead to misleading high false correlation between the pair. The 46 channel pairs that have this characteristic include F8-F4:F4-C4, F8-F4:F4-FZ, F7-F3:F3-C3, F7-F3:F3-FZ, F4-C4:F4-FZ, F4-C4:T4-C4, F4-C4:C4-CZ, F4-C4:C4-P4, F3-C3:F3-FZ, F3-C3:T3-C3, F3-C3:C3-CZ, F3-C3:C3-P3, F4-FZ:F3-FZ, F4-FZ:FZ-CZ, F3-FZ:FZ-CZ, FZ-CZ:C4-CZ, FZ-CZ:C3-CZ, FZ-CZ:CZ-PZ, T4-C4:C4-CZ, T4-C4:C4-P4, T4-C4:T4-T6, T3-C3:C3-CZ, T3-C3:C3-P3, T3-C3:T3-T5, C4-CZ:C3-CZ, C4-CZ:CZ-PZ, C4-CZ:C4-P4, C3-CZ:CZ-PZ, C3-CZ:C3-P3, CZ-PZ:P4-PZ, CZ-PZ:P3-PZ, C4-P4:P4-PZ, C4-P4:P4-O2, C3-P3:P3-PZ, C3-P3:P3-O1, T4-T6:T6-O2, T3-T5:T5-O1, P4-PZ:P3-PZ, P4-PZ:P4-O2, P3-PZ:P3-O1, T6-O2:P4-O2, T6-O2:O1-O2, T5-O1:P3-O1, T5-O1:O1-O2, P4-O2:O1-O2, P3-O1:O1-O2. A total of 207 channel pairs is therefore analysed in this paper.

The total number of feature values for this work is based on 9 estimations (PLV, MI, Corrmax, corrMean, CorrLag, MaxCoh, MeanCoh, MaxWCoh and MeanWCoh) × 6 bands (Full, Delta, Theta, Alpha, Beta and Gamma) × 207 pairs × 528 samples = 5,901,984. To explore the statistically significant differences between EG vs HC, and EG vs NEAD, one-way analysis of variance (ANOVA) was employed to select channels for each band and the estimation was undertaken using $p < 0.00001$ and $p < 0.05$ for two classification tasks respectively. ANOVA tests the null hypothesis, i.e. means of the tested groups are equal and the *p*-value indicates the statistical significance. Rejection of the null hypothesis leads to the conclusion that the two groups are statistically different. The selection of the threshold *p* was based on previous studies (Orekhova *et al.*, 2014; Vecchio *et al.*, 2016).

The K-Nearest Neighbour (KNN) algorithm was applied to perform the classification of the selected features. During the initial development of classification solutions, different machine learning algorithms were analysed, such as Support Vector Machine (SVM), decision tree and KNN etc. KNN tended to present performance superiority in this study compared with other methods and was selected as the main classifier in this paper. During the development of KNN, different values of k were tested. Initially, k was equal to 1 and was gradually increased to find the optimal one. Finally, k was set to 15 as it showed the best classification accuracy. For higher values of k (17 to 40) the accuracy was not improving and for a big k (above 40), the accuracy was dropping. Euclidean distance was applied in the KNN classification because of its better interpretability and performance (Prasath *et al.*, 2017).

The dataset was divided into 10 subsets and then cross-validation was undertaken. For each iteration of the 10-fold cross-validation, different subsets are used for training and testing. In the first iteration, the first subset is used for testing and the remaining subsets are used for training. The second iteration uses the second subset for testing and so on. To obtain the final result, an average of 10 classification accuracies is computed. Each accuracy comes from a single iteration of k-fold cross-validation. The 10-fold cross-validation is performed for 5 times by reshuffling data to gather 5 accuracy results statistically for each classification task. In the 10-fold cross-validation, 475 samples were used for training and 53 samples were used for testing. To further evaluate the machine learning algorisms performance, accuracy (Accu), sensitivity (Sens) and specificity (Spec) were calculated, which are defined as:

$$\boldsymbol{Accu} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \times \mathbf{100\%} \tag{13}$$

$$\boldsymbol{Sens} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \times \mathbf{100\%} \tag{14}$$

$$\boldsymbol{Spec} = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}} \times \mathbf{100\%} \tag{15}$$

where TP = True Positive, FN = False Negative, TN = True Negative, FP = False Positive. Moreover, the receiver operating characteristic (ROC) curve, and the area under the ROC curve (AUC) (Pyrzowski *et al.*, 2015; Lotte *et al.*, 2018) were used to assess the goodness of classification to select appropriate machine learning methods, k value for KNN and features. Specifically, ROC is constructed from true positive rate (TPR = sensitivity) in the vertical axis and false positive rate (FPR = 1-specificity) in the horizontal axis (Blinowska *et al.*, 2017). Besides, both ANOVA and multiple comparisons were performed to statistically compare the results of different brain connectivity estimations from different bands.

## 3. RESULTS

The section aims to report the classification results based on a single feature selected from different connectivity features, bipolar pairs and frequency bands.

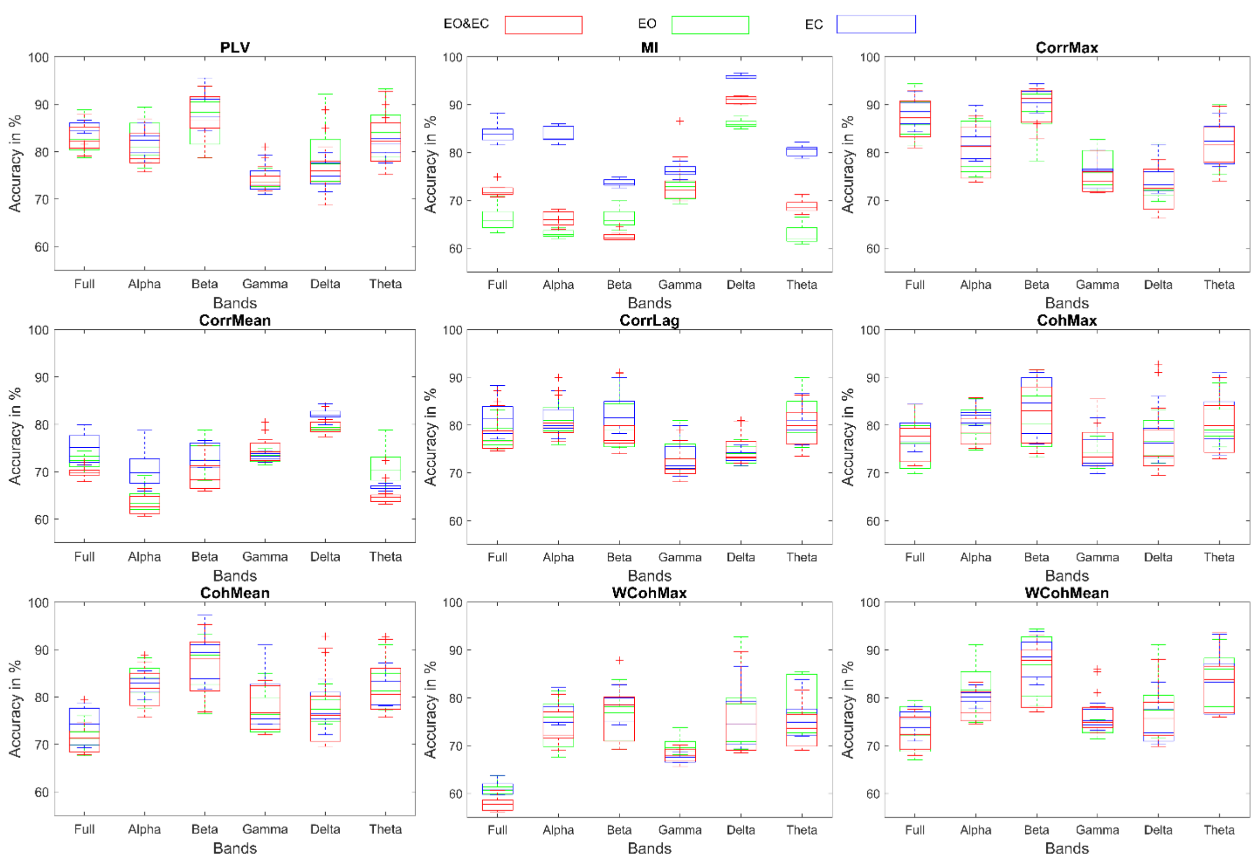
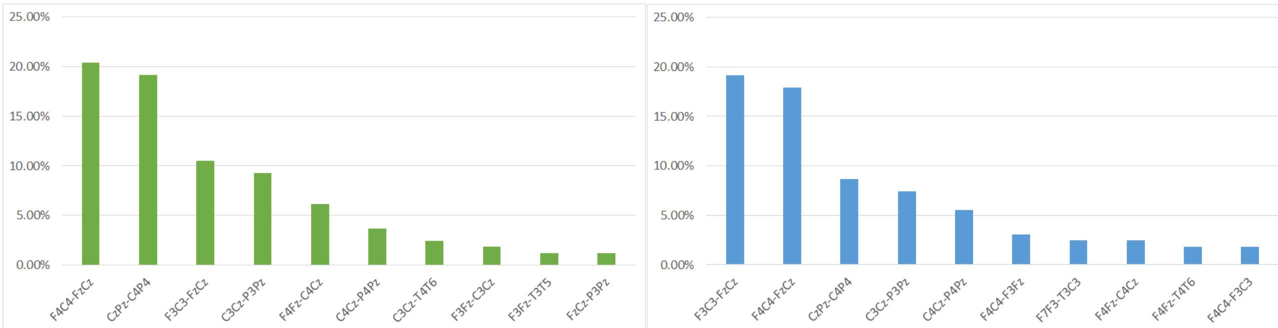



Fig. 4. The box charts of the classification accuracy of HC vs EG against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. Each box shows highest, lowest and median accuracy.




(a) (b)

Fig. 5. The top 10 bipolar pairs based on the percentage of appearance of top 3 classification performance between HC and EG for different connectivity measures and bands; (a) EO; (b) EC

### 3.1 Healthy Controls vs Epilepsy Group

Considering the considerable number of possible channel combinations (207), the one-way ANOVA for a threshold of $p < 0.00001$ was used to determine if the difference of each feature between HC and EG is statistically significant. Only the pairs with metrics that are significantly different between these two groups were selected for further classification. For example, for *mean coherence* in the Beta band, 67 out of 207 pairs showed a significant difference between these two groups. It is expected that statistically, the selected pairs will provide relatively high classification accuracy using the machine learning classifiers. As shown in Fig. S1 in Supplementary material, the pairs with a small *p*-value present a more distinguishable distribution of features than the ones with a relatively large *p*-value. However, it should be noted that the *p*-value cannot fully represent or replace the classification performance due to two reasons: (1) it focuses more on the mean of each group's features while machine learning classifiers pay more attention to the distribution of features, and (2) the cross-validation usually is used to determine the classification accuracy where testing data are not sampled, while to calculate the *p*-value all samples are used.

Fig. 4 plots the box charts of the classification accuracy of EG vs HC, against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. For the Theta band, all measures have similar performance with around 80% accuracy except *CorrMean* (<70%). For the Delta band, *MI* has exceptionally high performance with over 95% accuracy in EC while other bands have much lower accuracy (<80%). The rest 8 measures have only 70-80% accuracy. The Gamma band has relatively low accuracy across all measures (<80%). For the Beta band, *PLV*, *CorrMax*, *CohMean* and *WCohMean* have relatively good performance (>85%) while others are less than 80%. Overall, the Beta band has the best performance across all measures. The Alpha band has decent accuracy across all measures (around 80%). In terms of eye state, *MI* has the most distinguishable performance between EC and EO. Specifically, for all bands except the Beta band, the pattern of EC>(EC&EO)>EO can be observed. For other features' performance seems not to be significantly affected by the eye condition to an extent while combining EO and EC tends to slightly decrease the ability to discriminate EG from HC. This observation suggests that MI is more appropriate to classify these two groups for EC than EO and that the samples with different eye states should not be mixed.

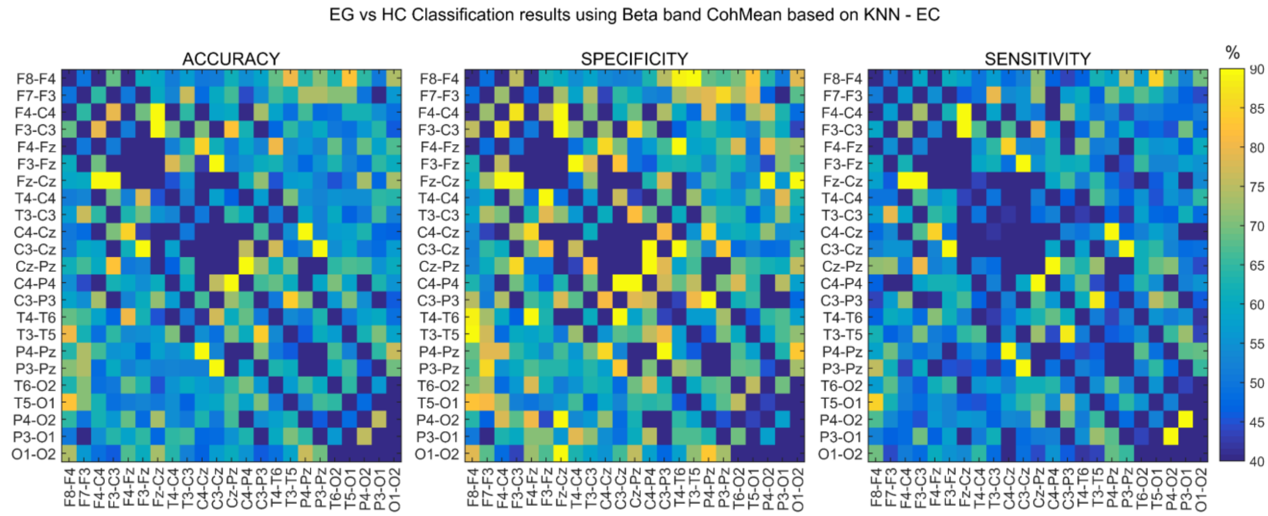
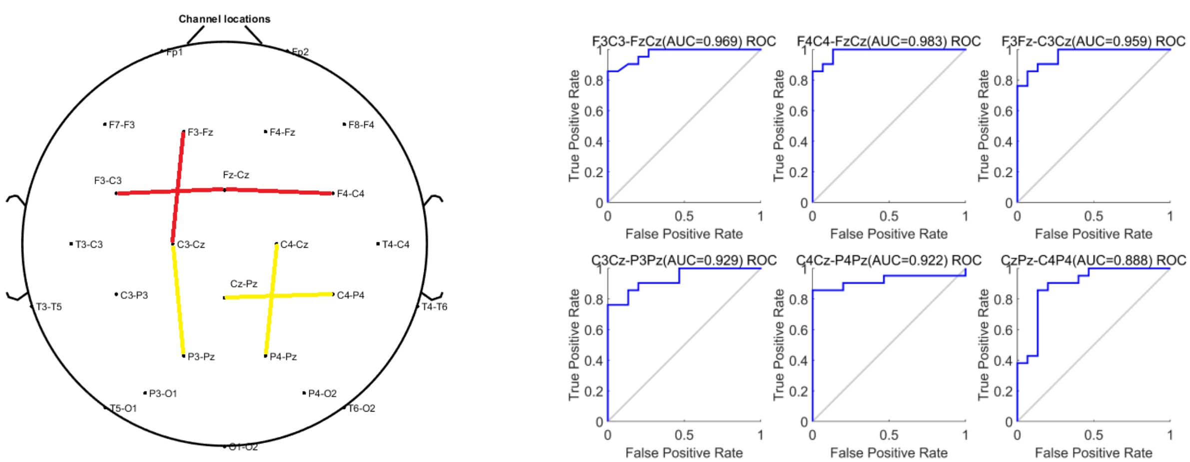



Fig. 6. The heatmap for the HC vs EG classification performance based on CohMean in Beta band for the EC state.

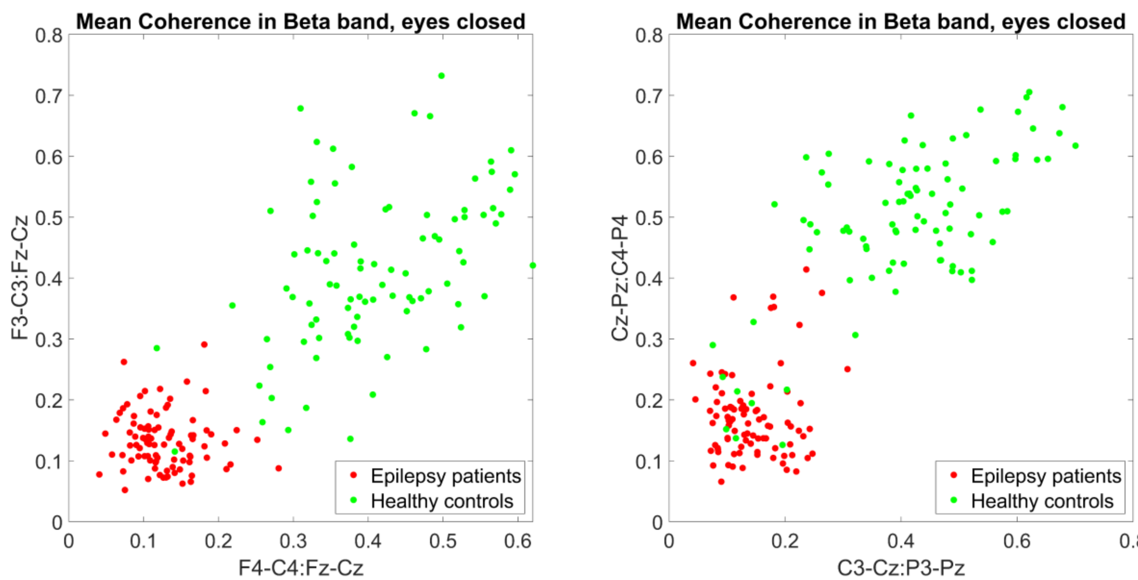



(a) (b)

Fig. 7. (a) The locations of 6 channel pairs corresponding to the highest classification accuracy (>90%) for the HC vs EG based on *CohMean* in the Beta band in the EC state (Red area is better than Yellow area); (b) The ROC and AUC of the 6 pairs.

To investigate which bipolar pairs consistently produced high classification accuracy, Fig. 5 plots the top 10 bipolar pairs based on the percentage of appearance of top 3 classification performance for different connectivity measures and bands. It can be observed that the top 3 pairs (F4C4-FZCZ, CZPZ-C4P4, F3C3-FZCZ) are identical for EO and EC. More detailed results using *CohMean* in the Beta band during the EC state are shown in Fig. 6. Besides, Fig. 7 represented the areas in a head map corresponding to the highest classification accuracy and the discrimination ability of identified features were evaluated by ROC and AUC.

To reveal the overall performance against different frequency bands, Fig. S2 in Supplementary material shows the top accuracy using the bipolar pair F4C4-FZCZ for different features and frequency bands. The result was achieved by averaging 5 trials. The best classification result (97.22%) was achieved using *CohMean* in the Beta band during the EC state for the bipolar pair F4C4-FZCZ. It can be observed that overall the Beta band has the best performance for EO, EC, and EO & EC, while the Gamma band has the worst performance consistently, which confirms the observations in Fig. 4. Besides, the evaluation of other machine learning methods and K selection refer to KNN supported by ROC and AUC, shown in Fig. S4 and Fig. S5 in Supplementary material. To further understand how the selected features contribute to the classification, Fig. 8 plots the clustering of *mean coherence* in the Beta band of F4C4-FZCZ, F3C3-FzCz, CzPz-C4P4 and C3Cz-P3Pz in the EC state, all of which have high classification accuracy (>90%). It can be observed that, for the HC subjects, the value of coherence is much higher than that of the EG subjects. Besides, it is also observed that the distribution of these features in EG is more concentrated, which may explain the classification sensitivity is always high, almost 100%, but with relatively low specificity. The machine learning has also been implemented using two features. The accuracy increases slightly, but there are still some HC samples misclassified as EG. The results are as shown in Fig. S6 in Supplementary material.




(a) (b)

Fig. 8. The distribution of mean coherence values in the Beta band of 4 channel pairs corresponding to the highest classification accuracy (>90%) for HC vs EG in the EC state.

### 3.2 Epilepsy Group vs NEAD

Fig. 9 plots the box charts of the classification accuracy of NEAD vs EG against different frequency bands using 9 distinct estimations. It can be observed that the overall accuracy is much lower than that of HC vs EG (about 20% less). Overall, the Gamma band produced a better performance (65-70%) than other 5 bands for all features. Especially for *PLV*, the Gamma band produces significantly higher accuracy which was witnessed by ANOVA and multiple comparisons ($p$<0.001).

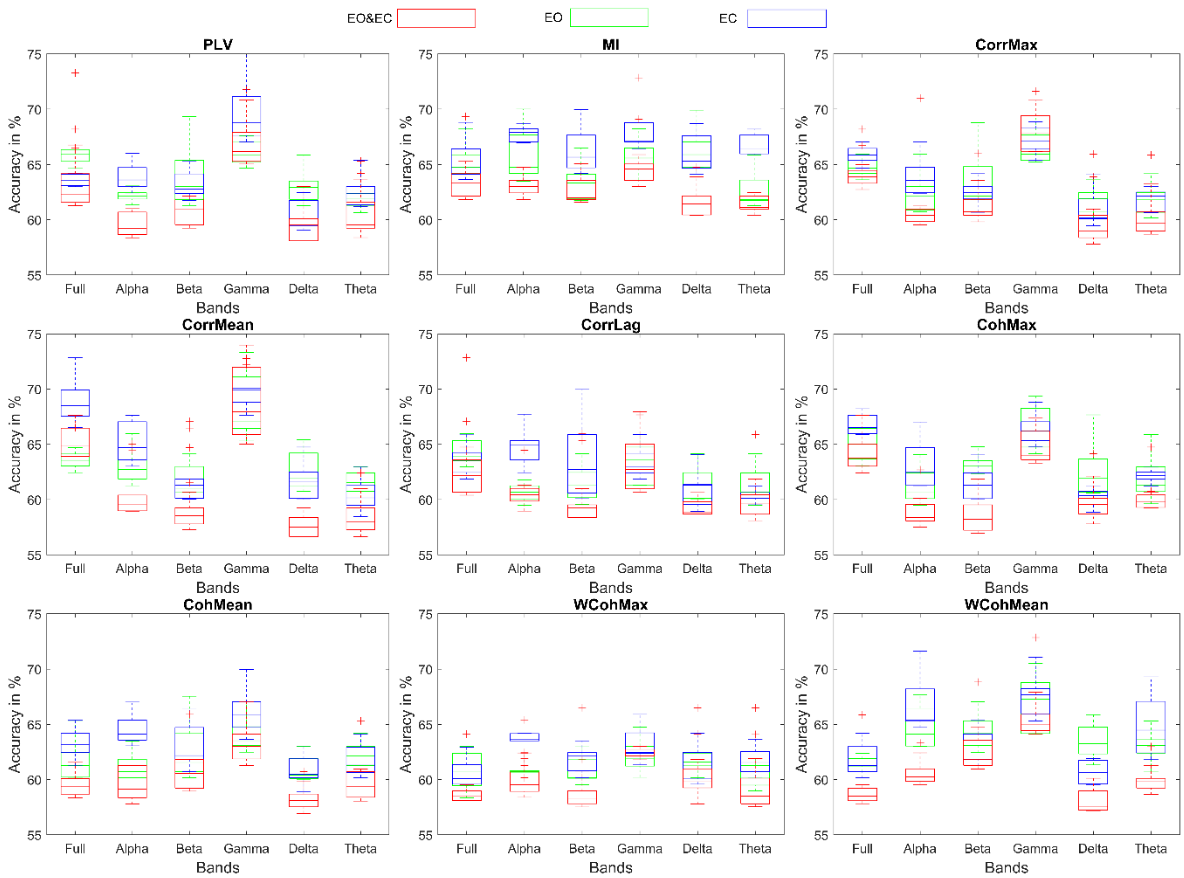



Fig. 9. The box charts of the classification accuracy of NEAD and EG against various frequency bands for the 9 selected measures, where only the top 10 pairs in terms of classification accuracy were considered. Each box shows highest, lowest and median accuracy.

Results (see Fig. S3) show that the best classification result (74.44%) was achieved using MI in the Delta band during the EO state for the bipolar pair T4T6-P4PZ. The second-best accuracy (74.24%) was achieved using PLV in the Gamma band during the EO state for the bipolar pair T3C3-CZPZ. Overall, the Gamma band has the best performance across all features and eye states. To further explore how the highest classification was produced, Fig. 11 presents the scatter plot of MI in the Delta band for T4T6-P4PZ and C4Cz-C3P3. The means of those two bipolar pairs for the EG is slightly lower than those of NEAD. ANOVA suggests that the means of both features for the two groups are significantly different ($p$<0.05).

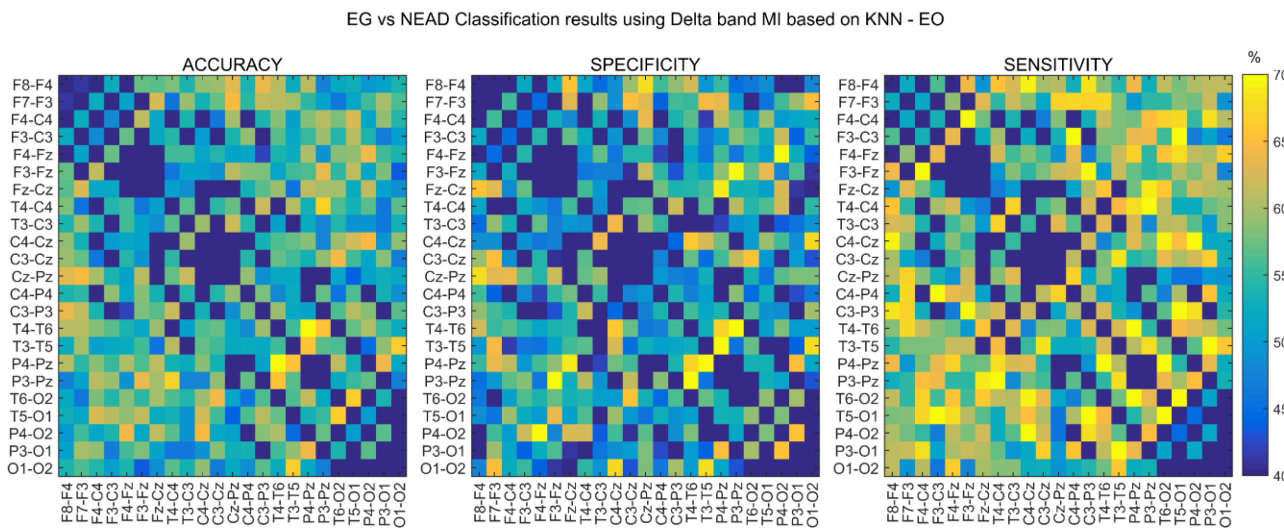



Fig. 10. The heatmap for the NEAD vs EG classification performance based on MI in Delta band for the EO state.

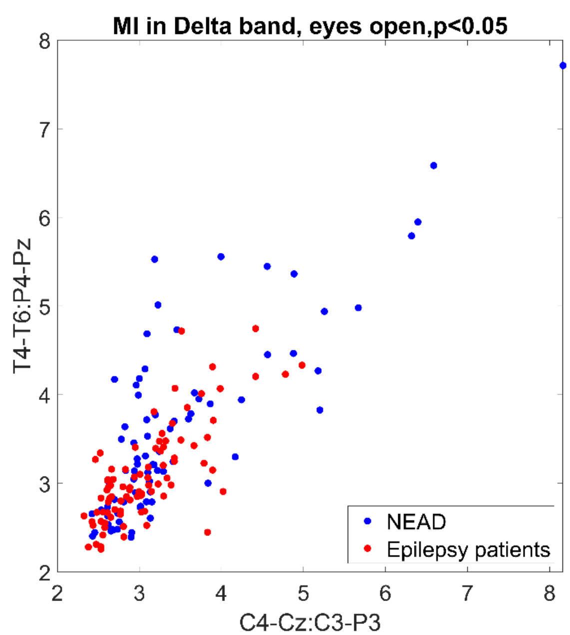



Figure 11 The distribution of MI values in Delta band of two channel pairs corresponding highest classification accuracy (>70%) for the NEAD vs EG for the EO state

## 4 DISCUSSIONS

This paper proposes a multivariate approach to investigate the potential of leveraging the association between two channels (i.e. EEG sensor level functional connectivity) to classify different groups. It should be noted that most of the state-of-the-art for this topic use the features extracted from single channels. Durongbhan *et al*. (2019) proposed a univariate method that uses the response of five frequency bands of each channel for classification. Similarly, a typical univariate method, power spectral density (PSD), was also applied to EEG recordings and it was found that theta (4-9 Hz) PSD ratio can contribute to evaluating the influence of neurofeedback training for epilepsy patients (Zhao *et al.*, 2009). Wan *et al.* (2019) suggested that alpha rhythm (8-12 Hz) PSD observed in EEG over human posterior cortex differs in distinct groups. Therefore, the univariate approach was applied to our database to critically compare the proposed and the existing method.

The spatial distributions of the frequency response of each channel are shown in Fig. S7-11 of Supplementary material, where the subjects are divided into 5 groups: EG with medication, EG without medication, NEAD with medication, NEAD without medication, HC. Fig. 12 shows the detailed spatial

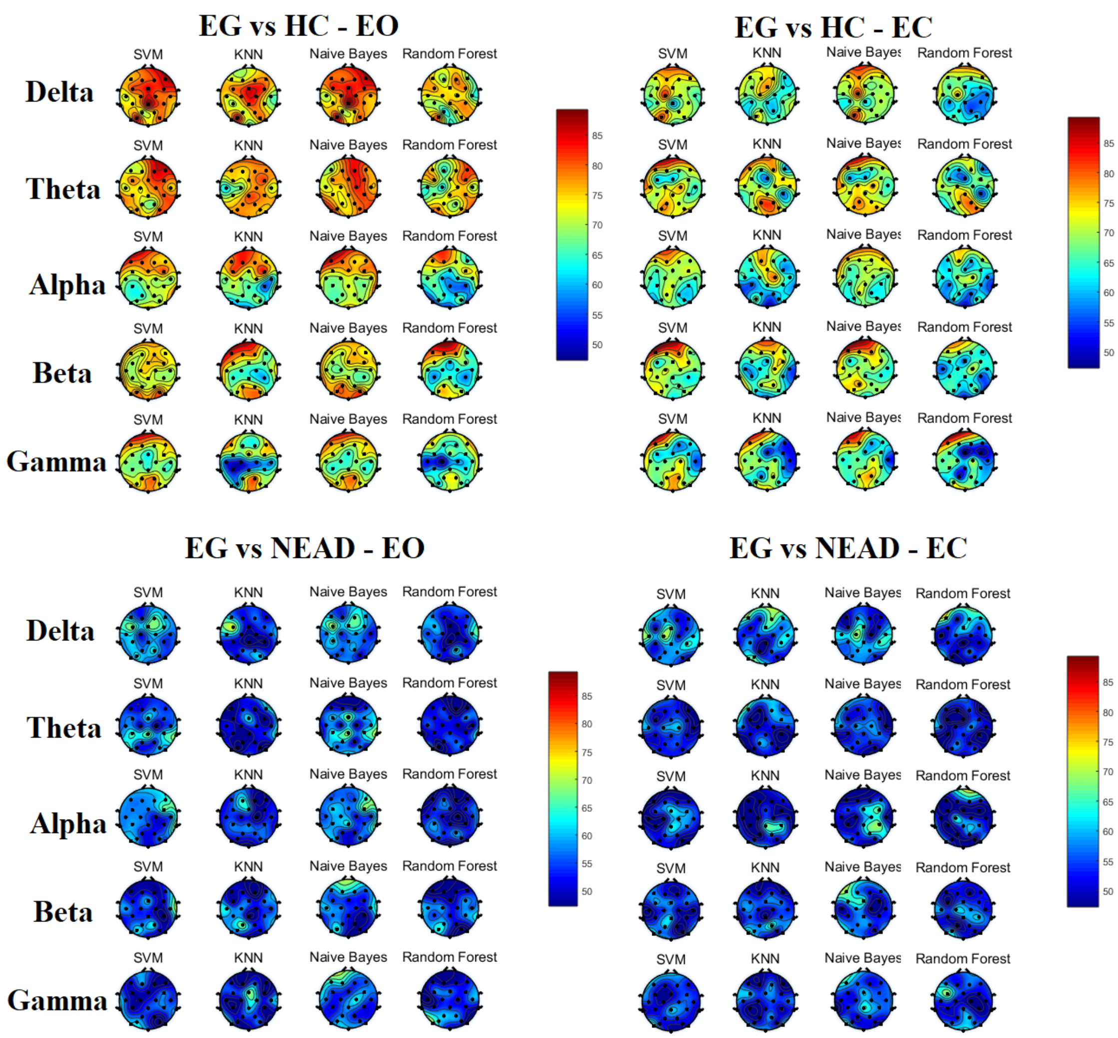



Fig. 12. The spatial distribution of classification accuracy using the univariate method, where the same colormap scale is used.

Table 2. The comparison of the highest classification using a single feature between the univariate approach and the proposed approach

| | Univariate approach | | The proposed approach | |
|---|---|---|---|---|
| | EO | EC | EO | EC |
| HC vs EG | 89.60% (Delta) | 82.80% (Delta) | 94.75% (Beta) | 97.22% (Beta) |
| NEAD vs EG | 72.20% (Theta) | 70.20% (Delta) | 74.44% (Delta) | 74.24% (Gamma) |

distribution of classification accuracy using the univariate method, where four classification methods were tested. It is observed that Naïve Bayes has a relatively good performance. Table 2 shows the comparison of the best classification result between the univariate approach (Naïve Bayes was used for classification) and the proposed multivariate approach for the same data. For HC vs EG, the proposed method based on multivariate analysis produced higher classification accuracy in comparison to the classic univariate method, particularly for EC, there is 15% improvement. It should be noted that the band with the best performance is different. For NEAD vs EG, the proposed method performs slightly better than the univariate approach for both EO and EC states.

## 5 CONCLUSIONS

In this work, we implement a framework comprising various sensor level functional brain connectivity features, on EEG epochs classified as normal on empirical inspection of interictal traces from a cohort of patients with generalised epilepsy. Despite no interictal abnormalities like IEDs - the empirical diagnostic hallmark of epilepsy - were included in this work, our framework achieved a very high classification accuracy between HC and EG cohorts, which outperforms the state-of-the-art single channel findings. Based on results from scalp EEG sensors, an accuracy of 97% was found for the functional connectivity estimates between the parasagittal frontal and midline regions for EC state and involved the beta frequency band. Unintuitively, as epilepsy in the ictal phase is typically characterised by the transient occurrence of abnormal, excessive or synchronous neuronal brain activity (Fisher *et al.*, 2014), our interictal findings show lower levels of synchronisation in this frontal region, in the beta band, in the Epilepsy patients in comparison to the findings from healthy controls (Fig. 8). This interictal deficit in the Beta band synchronisation in patients with generalised epilepsy was previously shown in a cohort of juvenile myoclonic epilepsy. This in part involved the frontal and parietal brain areas (Clemens *et al.*, 2013) that also consistently produced high classification accuracy in our work. However, our results have to be interpreted with caution as previous univariate EEG data analysis showed some spectral density differences in various phenotypes of generalised epilepsy syndromes (Clemens *et al.*, 2012), namely patients with absence seizures versus juvenile myoclonic epilepsy and epilepsy with generalised tonic-clonic seizures in isolation. The importance of this type of brain network connectivity approach was also shown from another angle in a study where decoupling of functional and structural connectivity, based on fMRI and tractography estimates was demonstrated for a cohort of patients with idiopathic generalised epilepsy in comparison to age-matched healthy controls (Zhang *et al.*, 2011). More work is required in well-characterised large cohorts of patients with different forms of generalised and focal epilepsies, ideally in drug naïve studies although clinical decisions render this latter requirement difficult to achieve.

Our findings demonstrate that seizure-free EEG recordings contain invisible information that can be revealed with appropriate methodology to identify patients with generalised epilepsy. The low sensitivity of a single EEG, at most at 55% (Pillai and Sperling, 2006), has been a long-standing diagnostic problem in clinical neurophysiology and translates to a high financial cost (several EEGs are frequently required to record IEDs) can result in time delay to treatment initiation and can also produce high levels of anxiety to patients, carers and relatives. Our work demonstrates that significant coherence data are confounded by power and phase effects, outside the observational capabilities of reporting physicians, can be successfully used to reveal spatiotemporal deficits in brain network organisation and behaviours that could be translated in clinical useful diagnostic tools. EEG epochs during periods of EO and EC differ significantly, reflecting the dynamic brain network changes between the two conditions. The discrepancy in the classification accuracy between EO and EC for HC vs the EG, based on MI, a technique able to capture nonlinear connectivity, demonstrates that advantages of different methods might have to be implemented in conjunction in future work for the development of a robust diagnostic framework. The results of the classification accuracy between EG and NEAD also demonstrates that there cannot be a one size fits all approach as the findings differ significantly when compared to the EG vs HC results. Specifically, the highest classification accuracies were achieved for synchronisation estimates within different frequency ranges (delta for EO and gamma for EC), involving different brain areas -the parietal and temporal regions - while additionally they were based on the two nonlinear methods (MI and PLV) implemented in this work. The roughly 20% lower classification accuracy (about 73%) between EG and NEAD is very reassuring, as not infrequently NEAD patients can also be experiencing epileptic seizures (Milán-Tomás *et al.*, 2018). This potential coexistence of the two conditions blurs the boundaries of “gold standard” labelling between the two, in keeping with the lower classification accuracy found in this work. Several EG and NEAD patients included in this work were receiving various AEDs which could influence the EEG recordings, extensively described in previous work (Höller, Helmstaedter and Lehnertz, 2018). However, 9 of our patients (5 from EG and 4 from NEAD) were on no medication. Therefore, it is highly unlikely that the remarkable classification accuracy of this study, particularly between HC and EG, was due to the effect of medication on EEG.

## REFERENCES


Akbarian, B. and Erfanian, A. (2020) 'Biomedical Signal Processing and Control A framework for seizure detection using effective connectivity , graph theory , and multi-level modular network', *Biomedical Signal Processing and Control*, 59, p. 101878. doi: 10.1016/j.bspc.2020.101878.

Amin, H. U., Yusoff, M. Z. and Ahmad, R. F. (2020) 'A novel approach based on wavelet analysis and arithmetic coding for automated detection and diagnosis of epileptic seizure in EEG signals using machine learning techniques', *Biomedical Signal Processing and Control*, 56. doi: 10.1016/j.bspc.2019.101707.

Battaglia, D. and Brovelli, A. (2019) 'Functional connectivity and neuronal dynamics : insights from computational methods', *The Cognitive Neuroscience*. Available at: https://hal.archives-ouvertes.fr/hal-02304918.

Bedo, N., Ribary, U. and Ward, L. M. (2020) 'Fast Dynamics of Cortical Functional and Effective Connectivity during Word Reading', 9(2), pp. 1–24. doi: 10.1371/journal.pone.0088940.

Blackburn, D. J. *et al.* (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 Sciences*, 8(7), pp. 1–19. doi: 10.3390/brainsci8070134.

Blinowska, K. J. *et al.* (2017) 'Functional and effective brain connectivity for discrimination between Alzheimer's patients and healthy individuals: A study on resting state EEG rhythms', *Clinical Neurophysiology*, 128(4), pp. 667–680. doi: 10.1016/j.clinph.2016.10.002.

Brodbeck, V. *et al.* (2012) 'EEG microstates of wakefulness and NREM sleep', *NeuroImage*, 62(3), pp. 2129–2139. doi: 10.1016/j.neuroimage.2012.05.060.

Brown, R. J. *et al.* (2011) 'Psychogenic nonepileptic seizures', *Epilepsy & Behavior*, 22(1), pp. 85–93. doi: 10.1016/j.yebeh.2011.02.016.

Clemens, B. *et al.* (2012) 'EEG-LORETA endophenotypes of the common idiopathic generalized epilepsy syndromes', *Epilepsy Research*, 99(3), pp. 281–292. doi: 10.1016/j.eplepsyres.2011.12.008.

Clemens, B. *et al.* (2013) 'Neurophysiology of juvenile myoclonic epilepsy: EEG-based network and graph analysis of the interictal and immediate preictal states', *Epilepsy Research*, 106(3), pp. 357–369. doi: 10.1016/j.eplepsyres.2013.06.017.

Cover, T. M. and Thomas, J. A. (2005) *Elements of Information Theory*, *Elements of Information Theory*. New York: JOHN WILEY and SONS, INC., 1991. ISBN: 0-471-20061-1. doi: 10.1002/047174882X.

Dash, G. K. *et al.* (2018) 'Interictal regional paroxysmal fast activity on scalp EEG is common in patients with underlying gliosis', *Clinical Neurophysiology*, 129(5), pp. 946–951. doi: 10.1016/j.clinph.2018.02.007.

Delgado-Restituto, M., Romaine, J. B. and Rodríguez-Vázquez, Á. (2019) 'Phase Synchronization Operator for On-Chip Brain Functional Connectivity Computation', *IEEE Transactions on Biomedical Circuits and Systems*, 13(5), pp. 957–970. doi: 10.1109/TBCAS.2019.2931799.

Dhiman, R., Saini, J. S. and Priyanka (2014) ‘Genetic algorithms tuned expert model for detection of epileptic seizures from EEG signatures’, *Applied Soft Computing Journal*, 19, pp. 8–17. doi: 10.1016/j.asoc.2014.01.029.

Durongbhan, P. *et al.* (2019) ‘A Dementia Classification Framework Using Frequency and Time-Frequency Features Based on EEG Signals’, *IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING*, 27(5). doi: 10.1109/TNSRE.2019.2909100.

Fani, M., Azemi, G. and Boashash, B. (2011) ‘EEG-based automatic epilepsy diagnosis using the instantaneous frequency with sub-band energies’, *7th International Workshop on Systems, Signal Processing and their Applications, WoSSPA 2011*, pp. 187–190. doi: 10.1109/WOSSPA.2011.5931447.

Fein, G. *et al.* (1988) ‘Common reference coherence data are confounded by power and phase effects’, *Electroencephalography and Clinical Neurophysiology*, 69(6), pp. 581–584. doi: https://doi.org/10.1016/0013-4694(88)90171-X.

Fisher, R. S. *et al.* (2014) ‘ILAE Official Report: A practical clinical definition of epilepsy’, *Epilepsia*, 55(4), pp. 475–482. doi: 10.1111/epi.12550.

Guevara Erra, R., Perez Velazquez, J. L. and Rosenblum, M. (2017) ‘Neural Synchronization from the Perspective of Non-linear Dynamics’, *Frontiers in Computational Neuroscience*, 11, p. 98. doi: 10.3389/fncom.2017.00098.

Höller, Y., Helmstaedter, C. and Lehnertz, K. (2018) ‘Quantitative Pharmaco-Electroencephalography in Antiepileptic Drug Research’, *CNS Drugs*, 32(9), pp. 839–848. doi: 10.1007/s40263-018-0557-x.

Horstmann, M. T. *et al.* (2010) ‘State dependent properties of epileptic brain networks: Comparative graph-theoretical analyses of simultaneously recorded EEG and MEG’, *Clinical Neurophysiology*, 121(2), pp. 172–185. doi: 10.1016/j.clinph.2009.10.013.

Kaya, Y. and Ertuğrul, Ö. F. (2018) ‘A stable feature extraction method in classification epileptic EEG signals’, *Australasian Physical and Engineering Sciences in Medicine*, 41(3), pp. 721–730. doi: 10.1007/s13246-018-0669-0.

Lachaux, J.-P. *et al.* (1999) ‘Measuring Phase Synchrony in Brain Signals’, *Human Brain Mapping*, pp. 194–208. doi: 10.1017/S0007680500048066.

Lioi, G. *et al.* (2017) ‘Directional connectivity in the EEG is able to discriminate wakefulness from NREM sleep’, *Physiological Measurement*, 38(9), pp. 1802–1820. doi: 10.1088/1361-6579/aa81b5.

Lopes, M. A. *et al.* (2019) ‘Revealing epilepsy type using a computational analysis of interictal EEG’, *Scientific Reports*, 9(1), pp. 1–10. doi: 10.1038/s41598-019-46633-7.

Lotte, F. *et al.* (2018) ‘A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update’, *Journal of Neural Engineering*, 15(3). doi: 10.1088/1741-2552/aab2f2.

Milán-Tomás, Á. *et al.* (2018) ‘An Overview of Psychogenic Non-Epileptic Seizures: Etiology, Diagnosis and Management’, *Canadian Journal of Neurological Sciences*, pp. 130–136. doi:

10.1017/cjn.2017.283.

Noachtar, S. and Rémi, J. (2009) 'The role of EEG in epilepsy: A critical review', *Epilepsy & Behavior*, 15(1), pp. 22–33. doi: 10.1016/j.yebeh.2009.02.035.

Nunez, P. L. *et al.* (1997) 'EEG coherency I: Statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales', *Electroencephalography and Clinical Neurophysiology*, 103(5), pp. 499–515. doi: 10.1016/S0013-4694(97)00066-7.

Orekhova, E. V. *et al.* (2014) 'EEG hyper-connectivity in high-risk infants is associated with later autism', *Journal of Neurodevelopmental Disorders*, 6(1), pp. 1–11. doi: 10.1186/1866-1955-6-40.

Pievani, M. *et al.* (2011) 'Functional network disruption in the degenerative dementias', *The Lancet Neurology*, 10(9), pp. 829–843. doi: 10.1016/S1474-4422(11)70158-2.

Pillai, J. and Sperling, M. R. (2006) 'Interictal EEG and the Diagnosis of Epilepsy', *Epilepsia*, 47(s1), pp. 14–22. doi: 10.1111/j.1528-1167.2006.00654.x.

Prasath, V. B. S. *et al.* (2017) 'Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier -- A Review', pp. 1–39. doi: 10.1089/big.2018.0175.

Pyrzowski, J. *et al.* (2015) 'Interval analysis of interictal EEG: Pathology of the alpha rhythm in focal epilepsy', *Scientific Reports*, 5, pp. 1–10. doi: 10.1038/srep16230.

Renzel, R. *et al.* (2017) 'Persistent generalized periodic discharges: A specific marker of fatal outcome in cerebral hypoxia', *Clinical Neurophysiology*, 128(1), pp. 147–152. doi: 10.1016/j.clinph.2016.10.091.

Rosch, R. *et al.* (2017) 'Network dynamics in the healthy and epileptic developing brain', *Network Neuroscience*, 2(1), pp. 41–59. doi: 10.1162/netn.

Sairamya, N. J. *et al.* (2018) 'An effective approach to classify epileptic EEG signal using local neighbor gradient pattern transformation methods', *Australasian Physical and Engineering Sciences in Medicine*, 41(4), pp. 1029–1046. doi: 10.1007/s13246-018-0697-9.

Sakkalis, V. *et al.* (2006) 'Assessment of linear and non-linear EEG synchronization measures for evaluating mild epileptic signal patterns', *Test*, 13(4), pp. 26–28.

Sakkalis, V. (2011) 'Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG', *Computers in Biology and Medicine*, 41(12), pp. 1110–1117. doi: 10.1016/j.compbiomed.2011.06.020.

Salman, H., Grover, J. and Shankar, T. (2018) 'Electroencephalographic Motor Imagery Brain Connectivity Analysis for BCI: A Review', 2733, pp. 2709–2733. doi: 10.1162/NECO.

Sarrigiannis, P. G. *et al.* (2014) 'Quantitative EEG analysis using error reduction ratio-causality test; validation on simulated and real EEG data', *Clinical Neurophysiology*, 125(1), pp. 32–46. doi: 10.1016/j.clinph.2013.06.012.

Sarrigiannis, P. G. *et al.* (2018) 'The cortical focus in childhood absence epilepsy; evidence from nonlinear analysis of scalp EEG recordings', *Clinical Neurophysiology*, 129(3), pp. 602–617. doi:

10.1016/j.clinph.2017.11.029.

Sheldon, R. J. G. and Agrawal, N. (2019) ‘Functional non-epileptic attacks: essential information for psychiatrists’, *BJPsych Bulletin*, 43(4), pp. 182–187. doi: 10.1192/bjb.2019.34.

Smith, S. J. M. (2005) ‘EEG in the diagnosis, classification, and management of patients with epilepsy’, *Journal of Neurology, Neurosurgery & Psychiatry*, 76(suppl_2), pp. ii2-ii7. doi: 10.1136/jnnp.2005.069245.

Tafreshi, T. F., Daliri, M. R. and Ghodousi, M. (2019) ‘Functional and effective connectivity based features of EEG signals for object recognition’, *Cognitive Neurodynamics*, 13(6), pp. 555–566. doi: 10.1007/s11571-019-09556-7.

Uhlhaas, P. J. and Singer, W. (2006) ‘Neural Synchrony in Brain Disorders: Relevance for Cognitive Dysfunctions and Pathophysiology’, *Neuron*, 52(1), pp. 155–168. doi: 10.1016/j.neuron.2006.09.020.

Varela, F. *et al.* (2001) ‘The brainweb: Phase synchronization and large-scale integration’, *Nature Reviews Neuroscience*, 2(4), pp. 229–239. doi: 10.1038/35067550.

Vecchio, F. *et al.* (2016) ‘Cortical connectivity and memory performance in cognitive decline: A study via graph theory from EEG data’, *Neuroscience*, 316, pp. 143–150. doi: 10.1016/j.neuroscience.2015.12.036.

Vijay Anand, S. and Shantha Selvakumari, R. (2019) ‘Noninvasive method of epileptic detection using DWT and generalized regression neural network’, *Soft Computing*, 23(8), pp. 2645–2653. doi: 10.1007/s00500-018-3630-y.

Visani, E. *et al.* (2010) ‘Photosensitive epilepsy: Spectral and coherence analyses of EEG using 14 Hz intermittent photic stimulation’, *Clinical Neurophysiology*, 121(3), pp. 318–324. doi: 10.1016/j.clinph.2009.12.003.

Wan, L. *et al.* (2019) ‘From eyes-closed to eyes-open: Role of cholinergic projections in EC-to-EO alpha reactivity revealed by combining EEG and MRI’, *Human Brain Mapping*, 40(2), pp. 566–577. doi: 10.1002/hbm.24395.

Watanabe, H. *et al.* (2018) ‘Effect of hyperventilation on seizures and EEG findings during routine EEG’, *Clinical Neurophysiology*, 129(5), p. e38. doi: 10.1016/j.clinph.2018.02.097.

Wendling, F. *et al.* (2010) ‘From intracerebral EEG signals to brain connectivity:Identification of epileptogenic networks in partial epilepsy’, *Frontiers in Systems Neuroscience*, 4, pp. 1–13. doi: 10.3389/fnsys.2010.00154.

Xie, S. and Krishnan, S. (2013) ‘Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis’, *Medical and Biological Engineering and Computing*, 51(1–2), pp. 49–60. doi: 10.1007/s11517-012-0967-8.

Zhang, Z. *et al.* (2011) ‘Altered functional–structural coupling of large-scale brain networks in idiopathic generalized epilepsy’, *Brain*, 134(10), pp. 2912–2928. doi: 10.1093/brain/awr223.

Zhao, L. *et al.* (2009) ‘Changes in EEG measurements in intractable epilepsy patients with neurofeedback training', *Progress in Natural Science*, 19(11), pp. 1509–1514. doi: 10.1016/j.pnsc.2009.03.010.

Zhao, Y. *et al.* (2018) 'A Wavelet-Based Correlation Analysis Framework to Study Cerebromuscular Activity in Essential Tremor', *Complexity*, 2018, pp. 1–15. doi: 10.1155/2018/7269494.

Zhu, G., Li, Y. and Wen, P. P. (2014) 'Epileptic seizure detection in EEGs signals using a fast weighted horizontal visibility algorithm', *Computer Methods and Programs in Biomedicine*, 115(2), pp. 64–75. doi: 10.1016/j.cmpb.2014.04.001.

**Cranfield University**

**CERES https://dspace.lib.cranfield.ac.uk**

School of Aerospace, Transport and Manufacturing (SATM) Staff publications (SATM)

2021-03-12

# Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity

Cao, Jun

Elsevier

Cao J, Grajcar K, Shan X, et al., (2021) Using interictal seizure-free EEG data to recognise patients with epilepsy based on machine learning of brain functional connectivity, Biomedical Signal Processing and Control, Volume 67, May 2021, Article number 102554

https://doi.org/10.1016/j.bspc.2021.102554

*Downloaded from Cranfield Library Services E-Repository*