IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 10, 2020, accepted December 27, 2020, date of publication January 5, 2021, date of current version January 14, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3049427

Automated Diagnosis of Major Depressive Disorder Using Brain Effective Connectivity and 3D Convolutional Neural Network

DANISH M. KHAN[1,2], (Student Member, IEEE), NORASHIKIN YAHYA[1], (Member, IEEE), NIDAL KAMEL[1], (Senior Member, IEEE), AND IBRAHIMA FAYE[1], (Senior Member, IEEE)

[1]Centre for Intelligent Signal and Imaging Research (CISIR), Electrical and Electronic Engineering Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar 32610, Malaysia

[2]Department of Electronic and Telecommunications Engineering, NED University of Engineering and Technology, Karachi 75270, Pakistan

Corresponding author: Norashikin Yahya (norashikin_yahya@utp.edu.my)

This work was supported in part by the Ministry of Education Malaysia under Higher Institutional Centre of Excellence (HICoE) Scheme awarded to Centre for Intelligent Signal and Imaging Research (CISIR) under Grant 015MA0-050(6), and in part by the Yayasan Universiti Teknologi PETRONAS under Grant YUTP-FRG 015LC0-031 and Grant YUTP-FRG 015LC0-292.

ABSTRACT Major depressive disorder (MDD), which is also known as unipolar depression, is one of the leading sources of functional frailty. MDD is mostly a chronic disorder that requires a long duration of treatment and clinical management. One of the critical issues in MDD treatment is the need for it's early diagnosis. Conventional tools in MDD diagnosis are based on questionnaires and other forms of psychiatric evaluations. However, the subjective nature of these tools may lead to misleading inferences. Recently, brain electroencephalography (EEG) signals have been used for the quantitative diagnosis of MDD. Nevertheless, a further improvement of the proposed methods in terms of accuracy and clinical utility is required. In this study, EEG signals from 30 MDD and 30 healthy control (HC) are used to estimate the effective connectivity within the brain default mode network (DMN). Then, effective connections between the major six regions of the DMN are used to train and test a three-dimensional (3D) convolutional neural network. Here, connectivity samples generated from half of the subjects are used for training while the rest are used for testing. The results show that the proposed MDD diagnosis algorithm achieved 100% accuracy,sensitivity and specificity in classifying MDD and HC test subjects.

INDEX TERMS 3D convolutional neural networks (CNN), brain effective connectivity, default mode network (DMN), major depressive disorder (MDD), partial directed coherence (PDC).

I. INTRODUCTION

Major depressive disorder (MDD), which is typically referred to as depression (unipolar), is a mental state which manifests itself via mood disorders, especially as low moods and aversions to activities. Based on the symptoms and severity, MDD can be classified as several different types such as bipolar disorder, atypical depression, dysthymia, psychotic depression, and postpartum depression [1]. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM–5), a person can be diagnosed as depressed if he/she shows low moods or loss of interest/pleasure for two weeks along with at least three more symptoms, such as an irregular appetite and weight fluctuation, slow thought processing, lethargy, feelings of worthlessness, loss of concentration and focus, and suicidal thoughts [2]. These symptoms may render the patient as functionally disabled. Thus, with approximately 7.5% of all Years Lived with Disability (YLD) [3], it is considered as a leading cause of disability worldwide, with more than 264 million people of all ages being directly affected, with more cases among women than men [4]. Owing to its heterogeneity and comorbid nature, the effective diagnosis of depression is a challenging task. The treatment of depression largely relies on its accurate and timely identification. Besides, it is highly probable that MDD is misdiagnosed as bipolar disorder (BPD) or other psychiatric disorders that are due to the common occurrence of depressive symptoms [5]. This misdiagnosis results in delayed or inappropriate medications which not only affect the efficacy of the treatment course but may also lead to further complications such as the development of drug resistance and deterioration in patients' mental condition. Hence, an accurate and timely diagnosis of

The associate editor coordinating the review of this manuscript and approving it for publication was Ludovico Minati.

MDD will give a better chance for effective treatment and lessen the risk of further damage. In addition, this error in prognostics may be due to the subjective nature of conventionally available diagnosing criteria of questionnaires and other human factors, including the expertise of the doctor, irregular patient history, and overstated depressive symptoms reported from the patient or the person attending him/her. This issue strongly suggests the incorporation of quantitative measures using brain imaging modalities such as magnetic resonance imaging (MRI), functional MRI (fMRI), and electroencephalography (EEG) for reliable diagnosis of MDD.

Over the past few years, the diagnosis of MDD has been performed using EEG-based machine-learning (ML) techniques owing to its non-invasive nature and high temporal resolution [6]–[9]. The automated EEG-based ML methods have proven their ability to correctly classify depressed and healthy subjects [7], [10]. Similarly, ML methods, especially support vector machines (SVMs) over fMRI-based features, have had valuable results in the identification of MDD patients. Recently, a depression diagnostic index has been proposed based on nonlinear features extracted from EEG data [11]. Radial basis function and linear discriminant analysis was performed in [12] with a classification accuracy of 93.33 % based on relative wavelet energy and wavelet entropy. Apart from conventional machine-learning techniques, few attempts have also been made to diagnose MDD using deep learning networks (DLNs). Acharya *et al.* [13] used a 13-layer convolutional neural network (CNN) over 15 healthy and 15 depressed subjects and obtained classification accuracies of 96% and 93.5% for the right and left hemispheres, respectively. A combination of 1D EEG data along with demographic information including gender and age are fed into a 1D CNN achieving 75.29% classification accuracy was also reported in [14]. In addition, [15] proposed a combined model of CNN and long short-term memory (LSTM) for one-dimensional EEG (1D-EEG) signals and achieved a classification accuracy of 98.32% for 33 depressed and 30 healthy controls. However, the combination of CNN and LSTM is a complex architecture which is computationally costly, and the classification is still not perfect.

Recently, brain connectivity has been analyzed in order to understand the effects of depression over brain networks. Aberrant brain connections have been found in depressed patients as compared to healthy subjects [16]–[18]. Olbrich *et al.* [19] demonstrated abnormal functional connectivity among different scalp locations in the frontal and temporal regions when compared with healthy controls. Xie *et al.* [20] used combination of brain functional network based on phase lag index (PLI) and a simple 2D CNN, giving 67.67% classification accuracy tested on 10 HC and 10 patients. Similar combination was also used by Li *et al.* [21] in which functional connectivity from different EEG bands was transformed into images and trained on 2-stacked CNN to achieve 80.74% classification accuracy for mild depression patients and HC. In [22], mixed feature matrices obtained from inter-hemispheric asymmetry and cross-correlation of EEG signals from 64 electrodes were used for training of a 2D CNN achieving 94.13% classification accuracy for 16 MDD and 16 HC. Although many attempts have been made to automatically diagnose MDD based on this functional synchronization, none of the proposed methods have achieved perfect diagnosis of MDD.

In this study, a new biomarker was developed to indicate the induced physiological changes in the human brain due to depression. These changes in the human brain are observed through the alternations in the exchanged causal effects between the brain's default mode network (DMN) regions owing to MDD. The causal effect of one region on another is referred to as brain effective connectivity. These effective connections, which are developed throughout the DMN regions, are fed into a three-dimensional (3D) CNN for the binary classification of the subject's data as MDD or healthy control. To the best of the authors' knowledge, this is the first attempt to diagnose MDD using a DMN's effective connectivity.

This paper is organized as follows: In section II, the resting state network (RSN) and the DMN components are described in detail, showing the regions involved in the DMN and the major functions of each one. The concept of effective connectivity is also briefly introduced under Section III. Next, Section IV describes the algorithm for estimating effective connectivity, known as partial directed coherence (PDC) algorithm. This section also covers the experimental setup, data acquisition, and code implementation, while Section V presents the results and discussion. Finally, Section VI concludes the paper and discusses on possible further work in the area.

II. DEFAULT MODE NETWORK

A subject is said to be in a resting state if he/she is not involved in any kind of activity or tasks. The brain networks which are active during this resting state are referred to as RSNs. Among these RSNs, a network of mesial prefrontal cortex, precuneus / posterior cingulate, and lateral parietal cortex is commonly referred to as a DMN, as shown in Fig. 1.

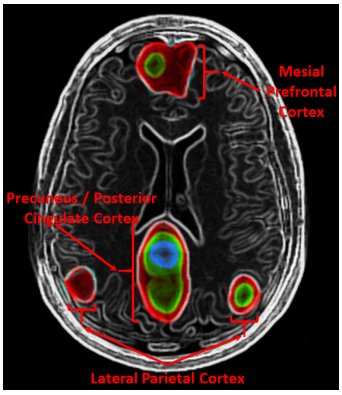


FIGURE 1. Axial view of the brain highlighting the DMN regions (reproduced from [23]).

DMN has recently emerged as an attractive research area for the scientific community as well as for clinicians because they exhibit a unique feature in that during the resting state, its activities intensify, and they become less active otherwise [24]–[26]. This less intensive activity of DMN rather than complete deactivation was also shown by Raichle and Snyder [27], where they observed the functional connectivity of mesial frontal gyrus and precuneus during the rest state as well as in tasks designed for the activation of working memory. This indicates that the DMN may facilitate or monitor the performance of active tasks, rather than simply getting deactivated [27].

Within the DMN network, precuneus has been given utmost importance owing to its strong correlation with other nodes of the same network in terms of connectivity. Besides, it acts as a mediator of intrinsic connections amongst them [28]–[30]. Owing to its intensive interconnections with other brain regions [30], [31], the involvement of precuneus has been observed in a variety of higher-order cognitive functions such as visio-spatial imagery, episodic memory retrieval, and self-processing. Similarly, lateral parietal cortex has significant connections with the medial temporal lobe, which includes hippocampal and entorhinal regions, a medial parietal lobe along with the anterior cingulate cortex, and a dorsolateral prefrontal cortex. The lateral parietal cortex has shown its involvement in cognitive processes such as spatial cognition, social cognition, working memory, attention [32], [33], and motor intent [34], [35]. Likewise, the mesial prefrontal cortex is involved in higher cognitive functions responsible for the processing, representation, and integration of social and affective information, social cognition [36], and self-referential memory [37] owing to its reciprocal connections with amygdala (emotion processing), hippocampus (memory), and temporal cortex regions involved in higher-order sensing [38]–[40]. Therefore, a study of the brain connectivity amongst DMN may increase researchers and clinicians' understanding of different physiological phenomena of human behavior. Accordingly, the variations in these connections can be treated as a new biomarker for different neurological disorders.

III. EFFECTIVE CONNECTIVITY

A brain connection is formed when neurons of one brain region are synchronized with neurons of other brain regions and interact dynamically by adjusting their rhythms with each other [41]. This interaction can be analyzed through either functional or effective connectivity techniques. Although functional connectivity (FC) [42] efficiently describes the connections between different regions using cross-correlation values, it fails to give the direction of influence between two regions, i.e., it will show the functional connection between regions A and B, but it cannot differentiate the amount of information sent to B from A, and vice versa. Similarly, it is unable to eliminate the third-party effect, i.e., if another region C influences B through A, FC displays it as a connection between A and B.

These limitations of FC can be resolved if a causal interaction (effective connectivity (EC)) is considered instead of cross-correlation. EC reflects the causal influence which one neuronal system exerts over another, thereby providing the direction of influence as well [2]. This influence can be directly estimated through signals, called data-driven EC, or it can be based on models to specify causal links between different brain regions, which is referred to as model-driven EC. Granger-causality (GC) [42] and dynamic causal modelling (DCM) [43] are classified as data-driven and model-driven EC techniques, respectively. GC assumes that 'if a signal can be predicted by the past information from a second signal better than the past information from its own signal, then the second signal can be considered causal to the first signal' [42].

GC was originally developed to predict the causality of bi-variate signals in the time domain. However, its applicability was further extended to the frequency domain by Gweke [44], where he presented interaction amongst various EEG frequency bands. Moreover, Gweke's work enabled the analysis of the coupling between EEG frequency bands which hold utmost biomedical significance. Later, the concept of bivariate GC was further altered to incorporate multivariate signals [45], [46]. Currently, directed transfer function (DTF) [47] and partial directed coherence (PDC) [48] techniques are two commonly used variants of GC. In this paper, PDC has been utilized for estimating the EC within the DMN regions and described in the Section IV-A and Section IV-B.

IV. MATERIALS AND METHODS

The general methodology for development of the algorithm for automated diagnosis of MDD using brain EC from EEG signals and 3D CNN is illustrated in Fig. 2. The 19-channel EEG signals are first cleaned using automatic artefact removal technique as described in Section IV-E. Subsequently, M continuous segments of length 2-sec is extracted for each subject for the calculation of PDC. Since, the brain is highly dynamic in nature and it is necessary to select proper EEG segment length which is sufficient enough for the reliable calculation of MVAR modelling parameters and yet small enough to identify minor changes in brain dynamics. Generally, in EEG processing, it is a common practice to consider segments of length less than 4-sec in order to ensure stationary of EEG. [49], [50]. Thus, each 2-sec segment gives one PDC matrix,which collectively yields M matrices of EC over the 19 channels for 1 subject in this study. Specifically, the PDC connectivity calculation will output $(19 \times 19 \times 64)$-connectivity matrix and the subsequent DMN connectivity extraction will reduce the connectivity matrix to $6 \times 6 \times 64$. The subsequent stage involves two steps. The first is where all the samples from the subjects selected for training the 3D-CNN network, referred to as the training subjects whereas in the second step the samples from testing subjects are the input of this trained network to obtain the classification results. Classification performance was then evaluated using accuracy, sensitivity and specificity. Details of the method to

calculate the EC and the training and testing of the proposed 3D CNN are covered in the subsequent sections.

A. PARTIAL DIRECTED COHERENCE (PDC)

Partial directed coherence is a frequency domain method based on the multivariate autoregressive model (MVAR) and the principle of partial coherence. Suppose that a set of n simultaneously observed time series

$$Y(t) = [y_1(t), y_2(t), \ldots, y_n(t)]^T \tag{1}$$

is represented by an autoregressive model of order p as given in Equation (2)

$$Y(t) = \sum_{l=1}^{p} \mathbf{A}_l \mathbf{Y}(t-l) + \boldsymbol{\varepsilon}(t), \tag{2}$$

where

$$\mathbf{A}_l = \begin{bmatrix} a_{11}(l) & \cdots & a_{1n}(l) \\ \vdots & \ddots & \vdots \\ a_{n1}(l) & \cdots & a_{nn}(l) \end{bmatrix} \tag{3}$$

is the coefficient matrix at the time lag l, and $\varepsilon(t) = \left[\varepsilon_1(t) \cdots \varepsilon_n(t)\right]^T$ is the zero mean multivariate Gaussian white process. The MVAR coefficient a_{uv} represents the effect of $y_v(t-l)$ on $y_u(t)$.

If $\mathbf{A}(f)$ is the frequency domain equivalent of coefficient matrix $\mathbf{A}_l$, then the Fourier transform $a_{uv}(f)$ of the elements $a_{uv}(l)$ can be obtained as,

$$a_{uv}(f) = \sum_{l=1}^{p} a_{uv}(l)\, e^{-i\left(\frac{2\pi}{p}\right)lf}. \tag{4}$$

By subtracting $\mathbf{A}(f)$ from n-dimensional identity matrix I, we get $\bar{\mathbf{A}}(f)$ then the PDC from electrode v to electrode u, denoted by $\pi_{uv}(f)$ is

$$\pi_{uv}(f) = \frac{\bar{a}_{uv}(f)}{\sqrt{\bar{a}_v^H(f)\,\bar{a}_v(f)}}, \tag{5}$$

where $\bar{a}_{uv}(f)$ represents the uv^{th} elements of matrix $\bar{\mathbf{A}}(f)$, and $(.)^H$ indicates conjugate transpose. The $\pi_{uv}(f)$ in Equation (5), represents the strength of the influence of electrode v over u at frequency (f) with values ranges from 0 for no connectivity to 1 for maximum connectivity. The basic Matlab code for the calculation of PDC is available at [51].

B. EFFECTIVE CONNECTIVITY IN DEFAULT MODE NETWORK

Selection of 6 electrodes that constitutes DMN is based on Brodmann areas (BA) specified for the brain regions as shown in Table 1. Here, the PDC matrix obtained shows connectivity of all the 19 channels over 64 bins. These $19 \times 19 \times 64$ matrices were then reduced to only six electrodes, which form the DMN, i.e., F3, F4, P3, P4, Fz, and Pz giving reduced matrix of size $6 \times 6 \times 64$ as illustrated in Fig. 2. Details of the algorithm for extraction of DMN connectivity is provided in Algorithm 1. The input to the algorithm is a PDC matrix obtained from 19 electrodes effective connectivity estimation followed by extraction of connectivities that mapped to DMN electrodes. This reduction from $19 \times 19 \times 64$ to $6 \times 6 \times 64$ is performed in order to eliminate all indirect causal effect of non-DMN regions over DMN regions. It should be noted here that the values of the total PDC matrix M varies from one subject, depending on the length of EEG after artefact removal.

TABLE 1. Mapping of EEG electrodes over DMN regions using Brodmann areas.

Electrode	DMN region	Brodmann Area
F3	Mesial Prefrontal Cortex	BA08/09L
F4	Mesial Prefrontal Cortex	BA08/09R
P3	Lateral Parietal Cortex	BA39/40L
P4	Lateral Parietal Cortex	BA39/40R
Fz	Mesial Prefrontal Cortex	BA08/09
Pz	Precuneus	BA07
L = left hemisphere , R= right hemisphere		

Algorithm 1 Extraction of DMN Connectivity

```
PDC_DMN ← [0]_{6×6×64}
PDC_ALL ← [PDC]_{19×19×64}
ALL_EL ← combination of 19 electrodes
DMN_EL ← Electrodes ↦ DMN regions
for all Rows(r) of PDC_ALL do
  for all Columns(c) of PDC_ALL do
    if ALL_EL(r, c) ∈ DMN_EL then
      PDC_DMN(r1, c1) ← PDC_ALL(r, c)
      r1 ← r1 + 1
      c1 ← c1 + 1
    else
      next iteration
    end if
  end for
end for
```

C. CONVOLUTIONAL NEURAL NETWORK

The 3D CNN will take in the $6 \times 6 \times 64$ PDC matrices as its input, which represent the connectivity of the EEG signals. The PDC matrices based on Equation (5) are calculated over six DMN channels at every (40/64)-Hz frequency bin, i.e., $f = 0.625b$, where $b = 1, 2, \ldots, 64$. Given the 3D PDC input, a 3D CNN will be employed for the classification of MDD versus healthy control (HC) from the EEG signal. The general architecture of our proposed 3D CNN is shown in Fig. 3, and consists of three convolutional layers, three batch normalization (BN), and three rectified linear unit (ReLU) activation layers, a global average pooling layer, a dropout layer and one fully connected layer. Each convolution layer is followed by a nonlinear activation function (ReLU) [52]. It should be noted that classification from the fully connected layer is based on the binary softmax regression.

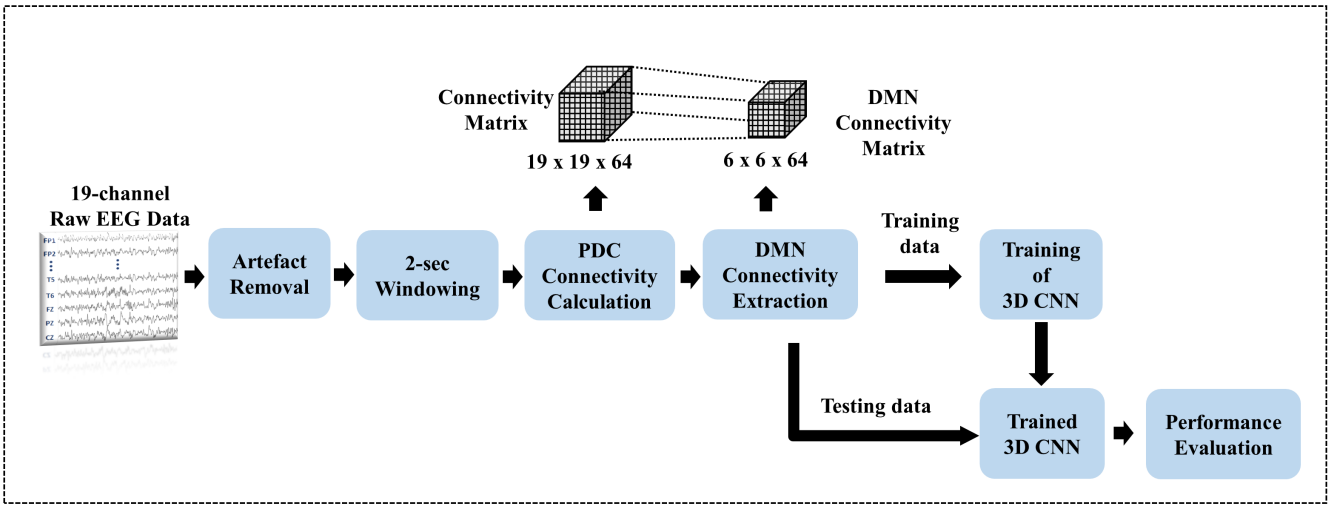


FIGURE 2. General pipeline for development of classification algorithm of MDD vs. HC using DMN effective connectivity and 3D CNN.

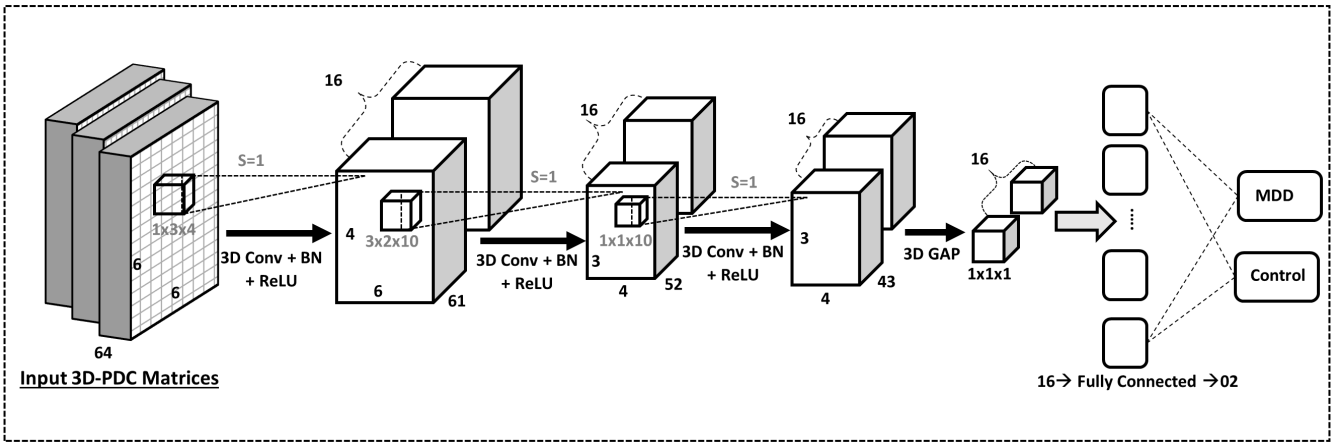


FIGURE 3. 3D CNN architecture for classification of MDD vs. HC using PDC matrices. Channel dimensions are in grey color, while 3D dimensions are in black. S = Stride, Conv = Convolution, BN = Batch normalization layer, ReLU = ReLU activation layer, 3D GAP = 3D global average pooling layer.

There are sixteen filters for each of the convolution layers, while the dimensions of the convolution kernels are $1 \times 3 \times 4$, $3 \times 2 \times 10$ and $1 \times 1 \times 10$, respectively, as shown in Fig. 3. The size of the convolution kernels is in decreasing order because no padding is applied; therefore, the kernel size needs to be adjusted for every convolution layer. Furthermore, random noise is added to the activation of each layer in order to reduce the issue of internal co-variance shift, producing improvement in training speed and reducing chances of over-fitting by using a batch normalization layer. The ReLU activation function is applied to the output of each batch normalization layer to map it to real numbers within a specific range in order to determine whether the information within the node is useful, i.e., whether or not to activate the node. The topological details of the proposed network are presented in Table 2.

The proposed network is trained using a mini-batch size of 64. The Adam optimizer is used as a solver for the training network, with a default initial learning rate of 0.001. All the other training options are also kept at their default values, except for L2 regularization, which is set to 0.0005 and 20% dropout in order to prevent over-fitting of data. The cross-entropy loss function is used for the final classification following a softmax layer, which assigns each input to one of the K mutually exclusive classes. The loss function for N number of samples is

$$loss = -\sum_{i=1}^{N}\sum_{j=1}^{K} t_{ij} \ln y_{ij} , \tag{6}$$

where t_{ij} indicates that the i^{th} sample belongs to the j^{th} class, and y_{ij} is the output for sample i for class j, i.e., the value from the softmax function. Training is performed by utilizing a GPU (Nvidia Quadro K620) for 60 epochs. The total training time is only about 6.1 min. This fast training time also makes the proposed 3D CNN framework a viable option for other neuroscience applications.

D. STUDY PARTICIPANTS

In this study, eye-close resting state EEG data were acquired from 30 MDD patients (Male: Female = 17: 13) with an

TABLE 2. Network architecture of the proposed 3D CNN along with the configuration and the corresponding number of trainable parameters and features.

Layer	Size	Trainable Parameters
3D image input	Input Size = 6 × 6 × 64 × 1	N/A
	No of features (6 × 6 × 64 × 1) = 2304	
Convolution	16 @ (1 × 3 × 4), Stride 1, No Padding, BN, ReLU	Weights : 1 × 3 × 4 × 1 × 16 = 192
		Bias : 1× 1 × 1 × 16 = 16
		Total Parameters : 192 + 16 = 208
Convolution	16 @ (3 × 2 × 10), Stride 1, No Padding, BN, ReLU	Weights : 3 × 2 × 10 × 1 6× 16 = 15,360
		Bias : 1 × 1 × 1 × 16 = 16
		Total Parameters : 15360 + 16 = 15,376
Convolution	16 @ (1 × 1 × 10), Stride 1, No Padding, BN, ReLU	Weights : 1 × 1 × 10 × 16× 16 = 2,560
		Bias : 1 × 1 × 1 × 16 = 16
		Total Parameters : 2560 + 16 = 2,576
Global Average Layer	Output = 1 × 1 × 1 × 16	N/A
Dropout	With probability 0.2	N/A
Fully Connected	Input = 16, Output = 2	Weights : 2 × 16 = 32
		Bias : 2 × 1 = 2
		Total Parameters : 32 + 2 = 34
Softmax	Activation :1 × 1 × 1× 2	N/A
Classification	Cross entropy	Total Parameters: 208 + 15376 + 2576 + 34 = 18,194

average age of 40 and a standard deviation of ±12.4. For HC, data from an age-matching group of 30 people (Male: Female = 19: 11) were acquired. The Diagnostic and Statistical Manual IV (DSM IV) was used to recruit 30 MDD patients who do not have any other psychotic symptoms. The recruitment was conducted at the Hospital Universiti Sains Malaysia (HUSM). Patients having a cumulative scores greater than 7 on Hospital Anxiety and Depression Scale (HADS) [53] were considered as MDD. Similarly, 30 HC subjects with no history of any psychiatric disorder were selected to participate in the study. Resting state EEG data used in this study is part of data used in [54] and is taken from CISIR's data repository.

The involvement of both the MDD patients and HC participants was voluntary. Their consent was obtained after a proper briefing about the experiment design, which was approved by the hospitals' ethics committee. To maintain consistency of the data, data acquisition for all the participants was performed at the same time of the day in a soundproof room. The participants were instructed to abstain from coffee, nicotine, and alcohol before EEG recordings.

E. EEG DATA ACQUISITION

A 5-min-long recording of EEG data was acquired from all of the participants during eye-close resting states. The data were recorded with a 19-channel EEG cap following the 10–20 electrode placement standard [55] with linked ear (LE) as the reference. The 19 electrodes covering the scalp include FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, FZ, CZ, and PZ. The 19-channel EEG cap was attached using an amplifier from Brain Master Systems (BrainMaster Discovery 24E). The data were recorded at a sampling rate of 256 Hz, and a bandpass filter (0.1–70 Hz) and a 50-Hz notch filter were applied. The EEGLAB software [56] was used for the automatic removal of artefacts.

Algorithm of artefact removal in EEGLAB was based on the artifact subspace reconstruction (ASR) method. This builtin plugin detects and removes muscles artefacts, eye blinks and sensor motions by comparing it with artefact free reference data as described in [57]. In this study, all the artefacts marked by the ASR method were removed and only the continuous 2-sec segments were used for the calculation of PDC.

The data used in this study and their findings will be made available for use upon reasonable request from the corresponding author after signing a formal data sharing and usage agreement. This study was approved by the ethics committee of the Hospital Universiti Sains Malaysia (HUSM).

F. EVALUATION OF 3D CNN

The generalization of the CNN was evaluated based on k-fold cross-validation (CV) set at $k = 10, 15$. The steps for k-fold CV are presented in Algorithm 2.

As shown in Algorithm 2, all the subjects from one class were randomly shuffled and stored in an array with fixed indices such that the same sequence of subjects could also

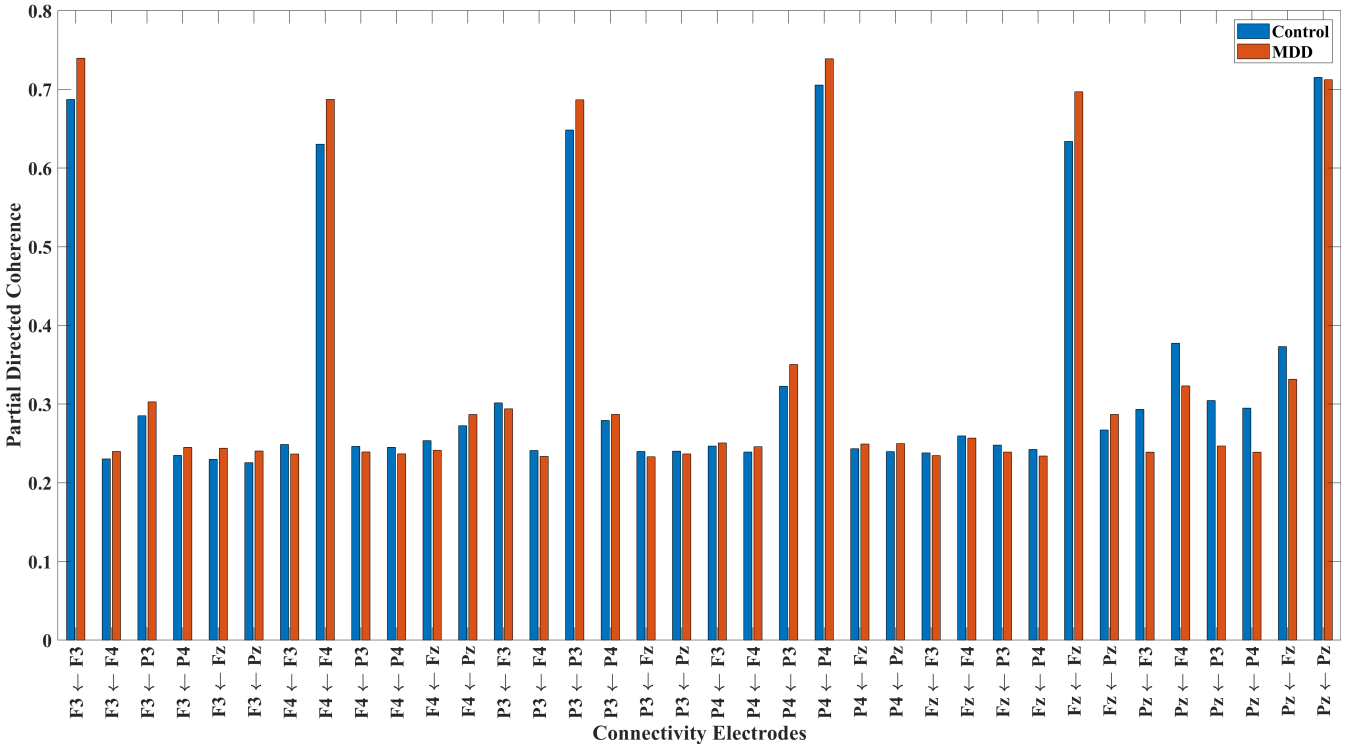


FIGURE 4. PDC connectivity of DMN electrodes for 30 MDD and 30 HC subjects.

Algorithm 2 k-Fold Cross Validation for CNN

$PDC_{MDD} \leftarrow [MDDSubjects]_{30}$
$PDC_{HC} \leftarrow [HCSubjects]_{30}$
$RandNo \leftarrow [Shuffle]_{30} \leftarrow only\ \ first\ \ time$
$RandGrp \leftarrow [RandNo]_{k-partitions}$

for $Iterations(j) = 1\ to 100$ **do**
  for all $FOLDS(kfld)$ **do**
    $Test_{MDD} \leftarrow RandGrp \in kfld$
    $Test_{HC} \leftarrow RandGrp \in kfld$
    $Train_{MDD} \leftarrow RandGrp \notin kfld$
    $Train_{HC} \leftarrow RandGrp \notin kfld$
    $CNN\ \ Training$
    $[Train_{MDD},\ \ Train_{HC}] \mapsto Network_{kfld}$
    $CNN\ \ Testing$
    $[Test_{MDD}, Test_{HC}] \mapsto ACC\_Net_{kfld}$
  end for
  $AVG\_ACC_{kj} \leftarrow Average(ACC\_Net)$
end for

$AVG\_ACC_{FOLD} \leftarrow Average(AVG\_ACC_{kj})$

be used for the next k-fold iteration. Likewise, the process was repeated for subjects in another class. Next, the subjects with corresponding indices from each class were arranged in unique groups according to the number of folds required for validation i.e (10 or 15). Subsequently, for every k-fold iteration, a group was selected from each class for testing while the remaining $k-1$ groups were used for training the 3D CNN. Thus, the outcome of each iteration was the classification accuracy of the testing group. After complete iterations, the accuracy values were averaged from 100 trials for the respective folds.

After k-fold CV and confirmation of generalization of the network, all the subjects were divided in a ratio of 0.5:0.5 for final training and testing. From the 30 MDD subjects, 15 were randomly selected for training, and the remaining 15 were used for testing. Similarly, the 30 HC subjects were also divided into training and testing classes. Accordingly, a total of 3965 PDC matrices were generated as the training set with ratio of MDD: HC = 1760: 2205 and 3902 for testing set with ratio of MDD: HC = 1910: 1992.

All the 3D-PDC training matrices were then fed to the 3D-CNN network evaluated earlier, as shown in Fig. 3, to obtain a trained network for classifying the unseen PDC matrices.

G. STATISTICAL ANALYSIS

To check the statistical significance of difference amongst MDD and HC subjects' individual connections, Multivariate Analysis of Variance (MANOVA) pairwise comparison was performed with significance level set at 0.05. Bonferroni correction was applied as an adjustment for multiple comparisons [58].

H. PERFORMANCE EVALUATION

Performance measures of how well the PDC connectivity can discriminate between MDD and HC subjects will be based

TABLE 3. *p*-values of 11 significant connections in DMN.

Connections	p-Value
F3 $\leftarrow F3$	0.001
F3 $\leftarrow FZ$	0.03
F4 $\leftarrow F3$	0.036
F4 $\leftarrow F4$	0
F4 $\leftarrow FZ$	0.048
P4 $\leftarrow P3$	0.028
P4 $\leftarrow P4$	0.031
FZ $\leftarrow FZ$	0
PZ $\leftarrow F3$	0.043
PZ $\leftarrow P3$	0.03
PZ $\leftarrow P4$	0.03

TABLE 4. Evaluation of the proposed classification algorithm using k-fold cross validation.

k	HC samples identification (%)	MDD samples identification (%)	Total (%)
10	94.5±14.73	94.67±4.99	94.95±7.32
15	95.02±13.98	95.68±4.76	95.65±7.62

on the performance of the developed 3D CNN. If *TP* is True Positive, *TN* is True Negative, *FP* is False Positive, and *FN* is False Negative, then the accuracy (*ACC*), sensitivity (*SEN*), and specificity (*SP*) are given as follows

$$ACC\ (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100, \tag{7}$$

$$SEN\ (\%) = \frac{TP}{TP + FN} \times 100, \tag{8}$$

$$SP\ (\%) = \frac{TN}{TN + FP} \times 100. \tag{9}$$

V. RESULT AND DISCUSSION

As described in section IV-E, *M* PDC matrices are initially calculated and averaged for all MDD and HC subjects in order to obtain a single PDC value for each class. The mean strengths of each connection within the DMN networks of MDD and HC are shown in Fig. 4.

By observing the connections having only significant differences in Fig. 4 obtained via MANOVA as described in IV-G and summarized in Table 3, it can be seen that in MDD patients, self-connectivity of electrodes is higher than that of HC. Eleven out of thirty-six connections have been found to be statistically significant with 95% confidence level as shown in Table 3. Bonferroni correction was applied on *p*-value for adjustment on multiple comparisons.

This high self-involvement of brain regions may be considered as an indicator of abnormal DMN in MDD patients. Apart from that, the network strengths for most of the significant connectivities in HC are relatively higher than that of MDD, which may reflect a degraded network formation between the frontal and parietal regions of MDD patients. The less active DMN connections in MDD patients may represent their executive dysfunction, suicidal thoughts, inability to focus, loss of interest, exhaustion, and lethargy; the discussion on which is beyond the scope of this paper. However, these changes may indicate that EC between different brain regions of DMN can be used to quantitatively identify MDD when compared to controls.

In order to evaluate the robustness and generalization of our proposed network, k-fold cross validation ($k = 10, 15$) technique was used. Average accuracies are provided in Table.4. High average accuracy, >95%, for 15-fold in Table 4 shows that the model is well generalized and DMN effective connectivity can be a potential biomarker for the diagnosis of MDD.

TABLE 5. Classification performance of 15 MDD (Subject 1 to 15) and 15 HC (Subject 16 to 30).

Test Subject No	PDC matrices (M)	Samples identified as		Accuracy (%)	Decision
		MDD	HC		
1	135	129	6	95.56	MDD
2	143	135	8	94.41	MDD
3	98	96	2	97.96	MDD
4	147	129	18	87.76	MDD
5	130	115	15	88.46	MDD
6	145	134	11	92.41	MDD
7	142	137	5	96.48	MDD
8	68	64	4	94.12	MDD
9	145	137	8	94.48	MDD
10	113	111	2	98.23	MDD
11	122	116	6	95.08	MDD
12	135	128	7	94.82	MDD
13	116	116	0	100	MDD
14	128	127	1	99.22	MDD
15	143	135	8	94.41	MDD
16	140	9	131	93.57	HC
17	129	0	129	100	HC
18	134	1	133	99.25	HC
19	149	1	148	99.33	HC
20	139	2	137	98.56	HC
21	123	0	123	100	HC
22	135	12	123	91.11	HC
23	143	4	139	97.20	HC
24	150	6	144	96	HC
25	147	0	147	100	HC
26	118	17	101	85.59	HC
27	136	7	129	94.85	HC
28	68	12	56	82.35	HC
29	157	0	157	100	HC
30	124	0	124	100	HC

TABLE 6. Comparison of the proposed technique with recent CNNs and deep learning techniques.

Ref	Year	Channels	Features	Input type	Classification method	No. of subjects		ACC (%)	SEN (%)	SPE (%)
						MDD	HC			
[22]	2020	64	Interhemispheric asymmetry & cross-correlation	2D Data	2D CNN	16	16	94.13	95.7	93.5
[59]	2020	19	EEG bands	2D Image	ResNet-50	46	46	90.2	N/A	N/A
			Spatial Locations		MobileNet			92.66		
[21]	2020	128	Functional Connectivity (Bands) and graph theory	2D Image	2-stacked CNN	24	24	80.74	N/A	N/A
[14]	2020	3	1D EEG data + Demographic attention	1D Data	1D CNN	81	89	75.29	66.2	83
[20]	2020	32	Phase Lag Index	2D Image	2D CNN	10	10	67.67	N/A	N/A
Proposed Technique		$19 \rightarrow 6$	Effective Connectivity-DMN	3D data	3D CNN	30	30	100	100	100

After evaluating the robustness of our proposed method, in the classification stage, as described in section IV-F, randomly selected M-PDC matrices of 15 MDD patients and 15 HC subjects were used to train the 3D CNN model, while the samples of the remaining 15 MDD and 15 HC subjects were used for blind testing purposes i.e. not a single matrix of testing was involved in data training. The results as a function of M have been presented in Table 5.

Here, it should be recalled that the variation in the total PDC matrices (M) for different subjects is primarily due to inconsistencies in the availability of clean 2-sec continuous segments between them. The results in Table 5 show the different values of M for 30 testing subjects, which vary for instance between 68 for subject 8 & 28 and 157 for subject 29. The subject-based decision is made depending on the total number of classification of PDC matrices for MDD and HC. If there are more PDC matrices classified as MDD, then the decision of the algorithm will be MDD, and vice-versa. The deep learning classification results in Table 5 show good statistical tendency towards the right decision for all of the subjects. This is indicated by the significant gaps between the number of correct and wrong classifications within the same subject.

The confusion matrix of the final subject-based classification process of 30 test subjects is shown in Fig. 5. From the confusion matrix, all MDD and HC cases are correctly classified with no misclassification, thereby, giving 100% accuracy sensitivity and specificity.

As machine learning and recently deep learning has found success in different brain disorders diagnosis, attempts are made to use EEG signals for classification of MDD using CNN or deep learning architectures. A comparison of the recently published studies based on EEG and CNN or deep learning models are given in Table 6.

Table 6 shows various recent deep learning methods, including 1D and 2D CNN networks that was used with different EEG features. Among the features used include EEG from multiple bands & spatial locations [59], interhemispheric asymmetry [22] and functional connectivity

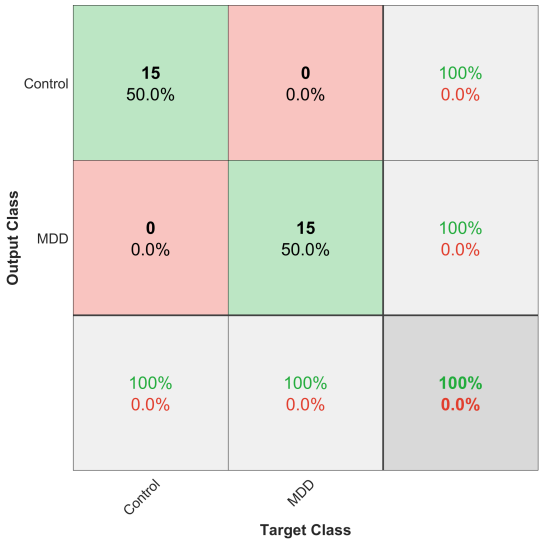


FIGURE 5. Confusion matrix for classification of 15-MDD and 15-HC test subjects using 3D CNN.

expressed in terms of PLI [20] and cross correlation [22]. Classification using functional connectivity [20], [22] as input of 2D CNN shows better accuracy in comparison to 1D CNN [14].

Our proposed technique uses brain effective connectivity as input to a 3D CNN gives 100% accurate classification between MDD and HC subjects. The technique employed causal effects amongst the DMN, which is a fundamental resting state network, may be treated as the biomarkers for MDD. Resting state diagnosis may eliminate the differences arising from goal directed tasks due to cognitive disability, age, education, gender, physical sloppiness, interest in task and sensori stimuli disability amongst subjects. In addition, resting state analysis may provide the actual on-going intrinsic activities inside the brain. In particular, resting-state connectivities between brain regions could provide an informative insight about the pathophysiology of MDD. The proposed MDD classification algorithm based on 6 DMN

electrodes as input to a 3D CNN is considered to be computationally fast. This indicates that the proposed technique has tremendous potential to be incorporated in clinical investigations of MDD with high degree of reliability, hence addressing the issue of subjectivity nature of questionnaire-based diagnosis.

VI. CONCLUSION

In this study, the differences between MDD and a HC are derived using the EC over the DMN regions. The results show a relatively higher level of information flow between the DMN regions with HC than for the MDD subjects apart from the abnormal high self connectivity of various DMN regions amongst MDD patients. These variations in the resting state connectivity strength of DMN may be related to the symptoms observed in MDD patients and require further investigation. Treatments like neurofeedback can be used in order to eliminate these abnormal variations of connectivity amongst DMN regions, which may help in faster recovery of MDD patients. Nevertheless, these changes indicate that the effective connections between the different brain DMN regions can be used to accurately classify the MDD subjects. To this end, a 3D CNN was developed for the classification of MDD and HC using DMN effective connections. This network was trained using 15 MDD and 15 HC subjects and was tested with the same number for both cases. The results show perfect accuracy in the classification of the MDD and HC subjects from their effective connections over the DMN regions. However, although it has perfect accuracy, the developed scheme requires further testing with more subjects before being employed in clinical applications.

DISCLOSURE STATEMENT

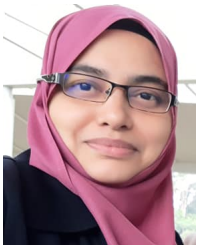
The authors declare no competing interests.

REFERENCES

[1] R. J. Baldessarini, "Psychiatric disorders," in *The Pharmacological Basis of Therapeutics*. New York, NY, USA: Macmillan, 1980, p. 391.
[2] A. P. Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Washington, DC, USA: American Psychiatric, 2013.
[3] *Depression and Other Common Mental Disorders: Global Health Estimates*, WHO, Geneva, Switzerland, 2017.
[4] WHO. (2019). *Depression*. Accessed: Jan. 20, 2020. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression
[5] I. C. Menezes, C. von Werne Baes, R. Lacchini, and M. F. Juruena, "Genetic biomarkers for differential diagnosis of major depressive disorder and bipolar disorder: A systematic and critical review," *Behav. Brain Res.*, vols. 357–358, pp. 29–38, Jan. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166432817303649
[6] W. Mumtaz, A. S. Malik, M. A. M. Yasin, and L. Xia, "Review on EEG and ERP predictive biomarkers for major depressive disorder," *Biomed. Signal Process. Control*, vol. 22, pp. 85–98, Sep. 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1746809415001287
[7] S. Mahato and S. Paul, "Electroencephalogram (EEG) signal analysis for diagnosis of major depressive disorder (MDD): A review," in *Nanoelectronics, Circuits and Communication Systems* (Lecture Notes in Electrical Engineering), vol. 511, V. Nath and J. Mandal, Eds. Singapore: Springer, 2019, doi: 10.1007/978-981-13-0776-8_30.
[8] S. Mahato and S. Paul, "Classification of depression patients and normal subjects based on electroencephalogram (EEG) signal using alpha power and theta asymmetry," *J. Med. Syst.*, vol. 44, no. 1, p. 28, Jan. 2020.
[9] P. Fernández-Palleiro, T. Rivera-Baltanás, D. Rodrigues-Amorim, S. Fernández-Gil, M. del Carmen Vallejo-Curto, M. Álvarez-Ariza, M. López, C. Rodriguez-Jamardo, J. L. Benavente, E. de las Heras, J. M. Olivares, and C. Spuch, "Brainwaves oscillations as a potential biomarker for major depression disorder risk," *Clin. EEG Neurosci.*, vol. 51, no. 1, pp. 3–9, Jan. 2020.
[10] D. Librenza-Garcia, B. J. Kotzian, J. Yang, B. Mwangi, B. Cao, L. N. P. Lima, M. B. Bermudez, M. V. Boeira, F. Kapczinski, and I. C. Passos, "The impact of machine learning techniques in the study of bipolar disorder: A systematic review," *Neurosci. Biobehav. Rev.*, vol. 80, pp. 538–554, Sep. 2017.
[11] U. R. Acharya, V. K. Sudarshan, H. Adeli, J. Santhosh, J. E. W. Koh, S. D. Puthankatti, and A. Adeli, "A novel depression diagnosis index using nonlinear features in EEG signals," *Eur. Neurol.*, vol. 74, nos. 1–2, pp. 79–83, 2015.
[12] S. Mahato and S. Paul, "Detection of major depressive disorder using linear and non-linear features from EEG signals," *Microsyst. Technol.*, vol. 25, no. 3, pp. 1065–1076, Mar. 2019.
[13] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. Adeli, and D. P. Subha, "Automated EEG-based screening of depression using deep convolutional neural network," *Comput. Methods Programs Biomed.*, vol. 161, pp. 103–113, Jul. 2018.
[14] X. Zhang, J. Li, K. Hou, B. Hu, J. Shen, J. Pan, and B. Hu, "EEG-based depression detection using convolutional neural network with demographic attention mechanism," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 128–133.
[15] W. Mumtaz and A. Qayyum, "A deep learning framework for automatic diagnosis of unipolar depression," *Int. J. Med. Informat.*, vol. 132, Dec. 2019, Art. no. 103983.
[16] A. E. Whitton, S. Deccy, M. L. Ironside, P. Kumar, M. Beltzer, and D. A. Pizzagalli, "EEG source functional connectivity reveals abnormal high-frequency communication among large-scale functional networks in depression," *Biol. Psychiatry. Cognit. Neurosci. Neuroimaging*, vol. 3, no. 1, p. 50, 2018.
[17] W. Mumtaz, S. S. A. Ali, M. A. M. Yasin, and A. S. Malik, "A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD)," *Med. Biol. Eng. Comput.*, vol. 56, no. 2, pp. 233–246, Feb. 2018.
[18] N. Fogelson, P. Diaz-Brage, L. Li, A. Peled, and E. Klein, "Functional connectivity abnormalities during processing of predictive stimuli in patients with major depressive disorder," *Brain Res.*, vol. 1727, Jan. 2020, Art. no. 146543.
[19] S. Olbrich, A. Tränkner, T. Chittka, U. Hegerl, and P. Schönknecht, "Functional connectivity in major depression: Increased phase synchronization between frontal cortical EEG-source estimates," *Psychiatry Res., Neuroimaging*, vol. 222, nos. 1–2, pp. 9–91, Apr. 2014.
[20] Y. Xie, B. Yang, X. Lu, M. Zheng, C. Fan, X. Bi, S. Zhou, and Y. Li, "Anxiety and depression diagnosis method based on brain networks and convolutional neural networks," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 1503–1506.
[21] X. Li, R. La, Y. Wang, B. Hu, and X. Zhang, "A deep learning approach for mild depression recognition based on functional connectivity using electroencephalography," *Frontiers Neurosci.*, vol. 14, p. 192, Apr. 2020.
[22] L. Duan, H. Duan, Y. Qiao, S. Sha, S. Qi, X. Zhang, J. Huang, X. Huang, and C. Wang, "Machine learning approaches for MDD detection and emotion decoding using EEG signals," *Frontiers Hum. Neurosci.*, vol. 14, p. 284, Sep. 2020.
[23] N. Challenged. (2015). *Know Your Brain: Default Mode Network*. Accessed: Jan. 13, 2020. [Online]. Available: https://www.neuroscientificallychallenged.com/blog/know-your-brain-defa%ult-mode-network
[24] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brain's default network: Anatomy, function, and relevance to disease," *Ann. New York Acad. Sci.*, vol. 1124, pp. 1–38, Mar. 2008.
[25] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: A network analysis of the default mode hypothesis," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 1, pp. 253–258, Jan. 2003.
[26] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proc. Nat. Acad. Sci. USA*, vol. 98, no. 2, pp. 676–682, 2001.
[27] M. E. Raichle and A. Z. Snyder, "A default mode of brain function: A brief history of an evolving idea," *NeuroImage*, vol. 37, no. 4, pp. 1083–1090, Oct. 2007.

[28] C. Rosazza and L. Minati, "Resting-state brain networks: Literature review and clinical applications," *Neurol. Sci.*, vol. 32, no. 5, pp. 773–785, Oct. 2011.
[29] M. Corbetta and G. L. Shulman, "Control of goal-directed and stimulus-driven attention in the brain," *Nature Rev. Neurosci.*, vol. 3, no. 3, pp. 201–215, 2002.
[30] M. Hampson, N. R. Driesen, P. Skudlarski, J. C. Gore, and R. T. Constable, "Brain connectivity related to working memory performance," *J. Neurosci.*, vol. 26, no. 51, pp. 13338–13343, Dec. 2006.
[31] P. Fransson and G. Marrelec, "The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis," *NeuroImage*, vol. 42, no. 3, pp. 1178–1184, Sep. 2008.
[32] A. E. Cavanna and M. R. Trimble, "The precuneus: A review of its functional anatomy and behavioural correlates," *Brain*, vol. 129, no. 3, pp. 564–583, Mar. 2006.
[33] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns, "Mapping the structural core of human cerebral cortex," *PLoS Biol.*, vol. 6, no. 7, p. e159, Jul. 2008.
[34] R. Cabeza, E. Ciaramelli, and M. Moscovitch, "Cognitive contributions of the ventral parietal cortex: An integrative theoretical account," *Trends Cognit. Sci.*, vol. 16, no. 6, pp. 338–352, Jun. 2012.
[35] G. F. Humphreys and M. A. L. Ralph, "Fusion and fission of cognitive functions in the human parietal cortex," *Cerebral Cortex*, vol. 25, no. 10, pp. 3547–3560, Oct. 2015.
[36] M. Desmurget, K. T. Reilly, N. Richard, A. Szathmari, C. Mottolese, and A. Sirigu, "Movement intention after parietal cortex stimulation in humans," *Science*, vol. 324, no. 5928, pp. 811–813, May 2009.
[37] S. Tumati, S. Martens, B. M. de Jong, and A. Aleman, "Lateral parietal cortex in the generation of behavior: Implications for apathy," *Prog. Neurobiol.*, vol. 175, pp. 20–34, Apr. 2019.
[38] S. W. Anderson, A. Bechara, H. Damasio, D. Tranel, and A. R. Damasio, "Impairment of social and moral behavior related to early damage in human prefrontal cortex," *Nature Neurosci.*, vol. 2, no. 11, pp. 1032–1037, 1999.
[39] M. Oscar-Berman, "Function and dysfunction of prefrontal brain circuitry in alcoholic Korsakoff's syndrome," *Neuropsychol. Rev.*, vol. 22, no. 2, pp. 154–169, Jun. 2012.
[40] T. Grossmann, "The role of medial prefrontal cortex in early social cognition," *Frontiers Hum. Neurosci.*, vol. 7, p. 340, Jul. 2013.
[41] J. N. Wood and J. Grafman, "Human prefrontal cortex: Processing and representational perspectives," *Nature Rev. Neurosci.*, vol. 4, no. 2, pp. 139–147, 2003.
[42] H. E. Wang, C. G. Bénar, P. P. Quilichini, K. J. Friston, V. K. Jirsa, and C. Bernard, "A systematic framework for functional connectivity measures," *Frontiers Neurosci.*, vol. 8, p. 405, Dec. 2014.
[43] K. Friston, R. Moran, and A. K. Seth, "Analysing connectivity with Granger causality and dynamic causal modelling," *Current Opinion Neurobiol.*, vol. 23, no. 2, pp. 172–178, Apr. 2013.
[44] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica, J. Econ. Soc.*, vol. 37, no. 3, pp. 424–438, 1969.
[45] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003.
[46] J. Geweke, "Measurement of linear dependence and feedback between multiple time series," *J. Amer. Stat. Assoc.*, vol. 77, no. 378, pp. 304–313, Jun. 1982.
[47] J. F. Geweke, "Measures of conditional linear dependence and feedback between time series," *J. Amer. Stat. Assoc.*, vol. 79, no. 388, pp. 907–915, Dec. 1984.
[48] Y. Hosoya, "Elimination of third-series effect and defining partial measures of causality," *J. Time Ser. Anal.*, vol. 22, no. 5, pp. 537–554, Sep. 2001.
[49] N. Nicolaou, S. Hourris, P. Alexandrou, and J. Georgiou, "EEG-based automatic classification of 'awake' versus 'anesthetized' state in general anesthesia using Granger causality," *PLoS ONE*, vol. 7, no. 3, Mar. 2012, Art. no. e33869.
[50] W. Hesse, E. Möller, M. Arnold, and B. Schack, "The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies," *J. Neurosci. Methods*, vol. 124, no. 1, pp. 27–44, Mar. 2003.
[51] A. Omidvarnia. (Aug. 2020). *Time-Varying EEG Connectivity: A Time-Frequency Approach*. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/33721-time-varying%-EEG-connectivity-a-time-frequency-approach
[52] M. J. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biol. Cybern.*, vol. 65, no. 3, pp. 203–210, Jul. 1991.
[53] R. P. Snaith, "The hospital anxiety and depression scale," *Health Qual. Life Outcomes*, vol. 1, no. 1, pp. 1–4, 2003.
[54] W. Mumtaz, L. Xia, M. A. M. Yasin, S. S. A. Ali, and A. S. Malik, "A wavelet-based technique to predict treatment outcome for major depressive disorder," *PLoS ONE*, vol. 12, no. 2, Feb. 2017, Art. no. e0171409.
[55] K. Sameshima and L. A. Baccalá, "Using partial directed coherence to describe neuronal ensemble interactions," *J. Neurosci. Methods*, vol. 94, no. 1, pp. 93–103, Dec. 1999.
[56] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004.
[57] T. R. Mullen, C. A. E. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-time neuroimaging and cognitive monitoring using wearable dry EEG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2553–2567, Nov. 2015.
[58] J. H. McDonald, *Handbook of Biological Statistics*, vol. 2. Baltimore, MD, USA: Sparky House, 2009.
[59] C. Uyulan, T. T. Ergüzel, H. Unubol, M. Cebi, G. H. Sayar, M. N. Asad, and N. Tarhan, "Major depressive disorder classification based on different convolutional neural network models: Deep learning approach," *Clin. EEG Neurosci.*, vol. 52, no. 1, pp. 38–51, 2020, doi: 10.1177/1550059420916634.

DANISH M. KHAN (Student Member, IEEE) received the bachelor's and master's degrees in telecommunications from the NED University of Engineering and Technology, Pakistan, in 2009 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Center for Intelligent Signal and Imaging Research (CISIR), Universiti Teknologi PETRONAS, Malaysia. Since March 2010, he has been a Faculty Member with the Department of Electronic and Telecommunications Engineering, NED University of Engineering and Technology. He is also a Research Scholar with CISIR, Universiti Teknologi PETRONAS. His research interests include signal processing, EEG, brain connectivity, seizure prediction, neurological disorders' diagnosis, and image processing.



NORASHIKIN YAHYA (Member, IEEE) received the B.Eng. degree (Hons.) in electronic engineering from The University of Sheffield, U.K., in 2001, the M.Sc. degree in electrical engineering from Lehigh University, USA, in 2004, and the Ph.D. degree in electrical engineering from the Universiti Teknologi PETRONAS (UTP), Malaysia, in 2015. She is currently a Senior Lecturer with UTP. She is also a member of the Centre for Intelligent Signal and Imaging Research (CISIR) Laboratory, one of the National Higher Institution Centre of Excellence (HICoE) status focusing on neuro signal and image analysis as its research niche area. Her current research interests include image segmentation and pattern recognition involving biomedical signals and images using deep learning architecture.

NIDAL KAMEL (Senior Member, IEEE) received the Ph.D. degree (Hons.) from the Technical University of Gdansk, Poland, in 1993. His Ph.D. work was focused on subspace-based array signal processing for direction of-arrival estimation. Since 1993, he has been involved in research projects related to estimation theory, noise reduction, optimal filtering, and pattern recognition. He is currently an Associate Professor with the Universiti Teknologi PETRONAS. His current research interests include brain signal and image processing using electroencephalography (EEG), fMRI, MRI, and fNIR for diagnosis and quantitative assessment of various brain disorders, including stress, anxiety, MDD, image processing for background initialization, visual periodic motion estimation, and magnification.

IBRAHIMA FAYE (Senior Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in mathematics from the University of Toulouse and the M.Sc. degree in engineering of medical and biotechnological data from Ecole Centrale Paris. He is currently an Associate Professor with the Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia. He is also with the Department of Fundamental and Applied Sciences, and the Centre for Intelligent Signal and Imaging Research (CISIR), a National Centre of Excellence. He has published over 150 papers in peer-reviewed journals and international conferences. He holds two patents in image processing. His research interests include machine learning, mathematics, signal and image processing, and science of learning. He is also the Vice-Chair of the IEEE Computational Intelligence Society (CIS), Malaysia.