IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 3, 2017, accepted July 14, 2017, date of publication July 26, 2017, date of current version August 14, 2017.
Digital Object Identifier 10.1109/ACCESS.2017.2731784

An EEG-Based Cognitive Load Assessment in Multimedia Learning Using Feature Extraction and Partial Directed Coherence

MOONA MAZHER, AZRINA ABD AZIZ, (Member, IEEE), AAMIR SAEED MALIK, AND HAFEEZ ULLAH AMIN
Centre for Intelligent Signal and Imaging Research, Department of Electrical and Electronic Engineering,
Universiti Teknologi PETRONAS, Seri Iskandar 32610, Malaysia

Corresponding author: Azrina Abd Aziz (azrina_aaziz@utp.edu.my)

This work was supported by the HiCoE Grant for CISIR, Ministry of Education, Malaysia, under Grant 0153CA-002.

ABSTRACT Assessing cognitive load during a learning phase is important, as it assists to understand the complexity of the learning task. It can help in balancing the cognitive load of postlearning and during the actual task. Here, we used electroencephalography (EEG) to assess cognitive load in multimedia learning task. EEG data were collected from 34 human participants at baseline and a multimedia learning state. The analysis was based on feature extraction and partial directed coherence (PDC). Results revealed that the EEG frequency bands and activated brain regions that contribute to cognitive load differed depending on the learning state. We concluded that cognitive load during multimedia learning can be assessed using feature extraction and measures of effective connectivity (PDC).

INDEX TERMS Cognitive load assessment, learning process, electroencephalography, and partial directed coherence.

I. INTRODUCTION

Because of the limited capacity of the human cognitive system for processing and holding information [1], acquiring information about user cognitive-processing capacity, memory workload, and task engagement [2] is important for researchers who study brain-computer interfaces (BCI), cognitive science, psychology, and human-computer interaction (HCI). Therefore, if cognitive activity exceeds the capacity of human working memory, performance might break down, resulting in failure to accomplish a cognitive task due to cognitive overload [3].

In the field of education, cognitive load assessment is often used to assess the productivity of learning materials, environment, and design, as a means to improve student learning. Measurement of cognitive load is based on Cognitive Load Theory (CLT), which analyzes both cognitive load assessment techniques and its methods and tools. Assessing cognitive load in this context helps users to maintain the optimal amount of cognitive load in different environments and high task demands [4]. CLT seeks to explain how and why some material is more difficult to learn than others. It is based on the proposition that the human brain uses two types of memory: short-term (working) and long-term (storage) memory. Short-term memory is seen as having a limited capacity, perhaps for as few as four ‘‘chunks’’ of information [5], and long-term memory is seen as having almost unlimited capacity [6]. Working memory is commonly defined as ‘‘a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning, and reasoning’’. Cognitive load refers to the load on working memory that a task demands. In recent years, researchers have increased focus on maintaining the best level of cognitive load during cognitive tasks [7]. Because the capacity and duration of cognitive processing is limited in the brain [5], any underloading or overloading of cognitive processing results in undesirable consequences such as degraded performance or failure to learn complex cognitive tasks or perform them correctly. Therefore, these situations must be avoided as much as possible, especially in critical decision-making applications, such as in air traffic control or military operations, in which the load requirements might be excessive and there is no room for error. Therefore, working memory architecture and its limitations should be a major consideration when considering cognitive-processing tasks [1]. Thus, the need to monitor and measure cognitive load is clear. In prior studies, researchers have used analytical and empirical methods to measure cognitive load [1], [8]. Analytical methods use techniques that rely on an expert's opinion, mathematical models, task analysis, and task specifications. Empirical methods gather the behavioral responses [9], [10] or physiological signals [10] of potential learners. Physiological techniques are justified by the assumption that changes in cognitive functioning are reflected by physiological variables such as heart rate variability [3], task-evoked brain activity [11], and skin conductance [12]. Changes to behavioral signals may be effected either through the changes in physiology or through different mental processing strategies. These measures can be used to visualize the detailed trends and patterns of cognitive loads such as instantaneous, peak, and accumulated load [13]. Because the brain is the seat of cognitive activity, cognition can be monitored with tools such as electroencephalography (EEG), which has been successfully used to monitor and measure different types of mental activities and workloads in cognitive science (e.g., relaxation, attention, and fatigue) and psychology (e.g., sleep, arousal, and anesthesia) [13].

EEG has also been shown to be an effective non-intrusive method for monitoring memory load, and it is highly sensitive to different cognitive states [14]. Analysis of EEG signals is used to diagnosis several brain disorders and cognitive processes. The EEG signal is a mixture of the desired information, which is related to the experimental conditions, and unwanted information (noise signal), which is generated by background brain processes. Analyzing EEG signals to assess cognitive load in learning is difficult when relying only on visual inspection or simple statistical parameters. Thus, adopting advanced computational techniques for processing EEG signals is necessary to extract the relevant information from raw EEG data in terms of cognitive load during learning. The basic steps required for extracting the desired information are preprocessing of raw EEG data, feature extraction, classification [15], and application of partial directed coherence.

Several feature extraction schemes exist, including time domain, frequency domain, and wavelet-transform. Among these feature extraction methods, wavelet transform works better for non-stationary EEG signals, making it highly effective. The wavelet features that are used for EEG analysis and clinical applications include (1) spectral features in which the commonly extracted features are wavelet coefficients [16] and wavelet entropy [17] and (2) statistical features in which the mean, median, and standard deviation are extracted. Past studies using EEG have shown good discrimination of cognitive tasks based on different workloads [18]. However, consistent discrimination for a given participant in a particular task with a constant amount of cognitive load remains lacking. Consequently, our study is based on wavelet-feature extraction with machine learning-classification techniques and partial directed coherence, using an EEG-based cognitive task with a persistent amount of workload. We focused on feature extraction using wavelet transform, and used the classification of the extracted features to categorize the EEG signals of learning and baseline mental states. Subsequently, we confirmed the findings with a coherence method based on EEG channels that uses partial directed coherence (PDC).

The objective of this paper was to analyze cognitive load by comparing baseline (eye closed and an eye opened states) and multimedia-learning EEG data using feature extraction and coherence analysis. Coherence measures the connectivity among different brain regions. We calculated the coherence of neighboring channels in the prefrontal, frontal, and parietal regions of the brain as a means to measure the effects of cognitive load during the given learning states.

Therefore, we performed experiments in which the learning state was repeated three times. The extracted features capable of discriminating cognitive load levels were empirically compared across different participants. We found that the descriptive statistics of different EEG-based features reflect the level of working load in the frontal regions of the brain. PDC results also depict the same response-like feature-extraction method. Further, this indicates that EEG is a promising real-time tool for measuring fairly rapid changes in cognitive load based on different learning states.

II. RELATED WORK

In educational psychology, EEG can be used to measure the neuronal response of changing levels of cognitive stimuli, making it is the most suitable measure for cognitive load assessment. EEG has been used for measuring the cognitive load of tasks and data analysis for over a decade. EEG signals comprise several frequency bands. The lowest frequency band is the delta band, ranging from 0.5 Hz to 4 Hz. It is followed by the theta band, ranging from 4 Hz to 7 Hz, and the alpha band, ranging from 7.5 Hz to 13.5 Hz. Alpha waves are the most prominent and dominant waves in the human EEG signal. Frequency ranges greater than 13 Hz are termed gamma and beta waves. Each range of EEG frequencies plays a different role in mental assessment.

Previously [19], have discussed how the power spectral density (PSD) of alpha waves remain synchronized during the resting state and become increasingly desynchronized with task difficulty. Another study has investigated cognitive load measurements, showing that alpha and theta waves play a great role in measuring cognitive load. Delta wave properties can be used to discriminate cognitive states [19]. Other researchers have proposed different feature-extraction methods for measuring cognitive load and for studying brain waves during cognitive tasks [20], [21]. The PSD of alpha waves is a widely used EEG feature that discriminates cognitive load in different brain states. Here, we hypothesized that low frequency waves (especially alpha waves) in the brain would be the primary signal to reflect cognitive processing (Karkare, Saha, & Bhattacharya, 2009). Classification techniques based on time, frequency, and wavelet-domain feature-extraction techniques have been reported in the literature for the last two decades [22], [23]. The techniques based on time, frequency, or a combination of both domains are used in different classification models. They provide an optimal set of features for further assessing the performance of different classifiers that attempt to deliver the highest classification performance. The features that are based on the time domain include permutation entropy and approximate entropy [24], as well as non-linear features such as the Hurst component [25], Hjorth parameters, Lyapunov exponent [26], and fractal dimension. The frequency-domain features are based on power or energy of the signal. Power features include absolute power, relative power, and the ratio between relative and absolute power [27]. Another class of features combines frequency and time features [28], includes the well-known Stockwell transform proposed by Hariharan *et al.* [29], which can be used to extract features from EEG signals and classify different cognitive tasks. They reported that the accuracy of their classification algorithm was between 84.72% and 98.95%. Noshadi *et al.* [30] reported empirical mode decomposition and implemented this model for time- and frequency-domain features in assessing cognitive tasks. They used linear classifiers (nearest neighbor and linear discriminant analysis) and reported 97.78% accuracy. Guo *et al.* (2013) implemented weighted Support Vector Machine (SVM) with safe features and reported 85.4%–97.5% accuracy in classifying cognitive tasks. Zhang *et al.* (2010) reported 72.4%–76.4% classification accuracy using power frequency features and a classifier (Fischer's discriminant) for EEG classification of cognitive tasks. Hosni *et al.* [31] introduced the radial basis function (RBF) as an SVM kernel, which provided an accuracy of 70%. Xue *et al.* [32] also proposed an RBF classification technique and reported 85.3% accuracy using cognitive tasks. Zhiwei and Minfen [33] reported 87.5%–93% accuracy with wavelet-based entropy as a feature and SVM as a classification technique. Lin and Hsieh [34] have used EEG-power features for cognitive tasks and reported 78.31% accuracy using a neural network classification technique. Rodrıguez-Bermudez *et al.* [35] applied wavelet-based features with 67.96%–80.71% accuracy using an SVM classification method. They used a linear kernel in an SVM classification experiment that included four participants. Karkare *et al.* [36] presented cognitive task-based classification with an artificial neural network (ANN) using a scaling exponent as a feature and reported a classification accuracy greater than 80%. Amin *et al.* [37] recorded eyes-opened and learning EEG states and used the discrete-wavelet transform-based feature extraction technique for classification of complex cognitive tasks. The relative wavelet energy was calculated as a feature from the EEG signals after applying the discrete wavelet transform. Four classifiers were used for the classification. The authors concluded that among all classifiers, SVM provided 98% accuracy. According to our knowledge, few studies have investigated learning-based cognitive tasks, and in those that have, the duration intervals for EEG recordings were short.

Here, the connectivity of different nodes within neuronal networks were measured and analyzed. There are many methods to measure the connectivity of brain regions based on EEG signals. Interpretation of connectivity measures from sensor-level recordings is not straightforward, as these recordings suffer from a low spatial resolution and are severely corrupted by the effects of field spread. To overcome these difficulties, several methods for connectivity analysis have been reported using the temporal dynamics of brain sources that were reconstructed from scalp EEG/MEG signals [38]. Two types of brain connectivity are explained in the literature: functional and effective. Functional connectivity is based on different time and frequency domain methods. These methods have their advantages and disadvantages [39]. Recently, partial directed coherence (PDC) has been generalized to time-varying and vector-based multivariate analyses [40]. Based on the methods of connectivity studies, PDC provides suitable explanation for EEG channels related to baseline activity as well as those related to learning.

III. MATERIALS AND METHODS

A. DESCRIPTION OF THE EEG DATASET

The dataset has been used previously [41]. The participant details, experimental tasks, experimental procedure, and ethics approval are briefly described here. Thirty-four healthy individuals (aged 20–30 years) participated in this study. None of the participants had no prior knowledge of the experimental tasks and all were instructed during the experiment to focus on the learning material. All participants had normal or corrected-to-normal vision. Ethical approval was obtained from UTP management, and all participants provided their informed consent before performing the experiment.

B. DESCRIPTION OF EEG RECORDINGS

EEG signals were recorded from a 128-channel EEG device (HydroCel Geodesic Sensor Net (Electrical Geodesic Inc., Eugene, OR, USA). The sampling frequency was 250 Hz and the impedance was kept below 50 KΩ. A band-pass filter was used to extract the frequencies between 0 and 100 Hz. ECG signal was recorded using EGI Polygraph input box (PIB) and EOG signal recorded using the specified electrodes in the 128 channels montage of EGI. The purpose of recordings EOG and ECG signals was to detect artifacts (eye blinks, eye movements, and heart beats) for pre-processing EEG data.

1) EYES-CLOSED TASK

This task was performed at the start of the experiment. EEG recordings were measured for 5 minutes for each participant. During this task, participants were asked to keep their eyes closed and try to relax. The eyes-closed (EC) recording was used as the baseline state.

2) LEARNING TASK

This task was performed after the EC recordings. During this task, multimedia animations based on plant anatomy were shown to each participant using Eureka Software (https://www.designmate.com/Product.aspx). All recordings were made in a controlled environment for 10 minutes. There were three learning-task recordings, termed L1, L2, and L3 (for details about the learning task, see [37]. L1 was the period when participants tried to learn the material in the multimedia animations for the first time, while L2 and L3 were the periods during which the same learning content was repeated a second and third time, respectively. These periods constituted the learning states.

3) MEMORY RECALL TASK

A memory recall test was used consisting of twenty multiple-choice questions (MCQs) covering the newly learned material in the learning task. Each MCQ comprised of a brief question statement with four options as possible correct answers. Participants were given thirty seconds to answer each MCQ within a maximum limit of 10 min total. They were asked to press a numeric key on the keyboard, serially numbered #1 to #4 corresponding to each possible answer.

C. METHODS

We used three types of classifiers (Naïve Bayes, Linear Kernel, and RBF Kernel) and three spectral features based on entropy for classification. PDC was applied to determine the connectivity patterns among the selected channels of the brain. Four distinct frequency bands were used to extract the spectral features: delta (0–3.5 Hz), theta (3.5–7.5 Hz), alpha (7.5–13.5 Hz), and beta (13.5–26 Hz). Raw EEG data has multiple artifacts that distort the signal information. Thus, pre-processing was performed on the raw EEG data to remove these artifacts. The complete analysis was based on the frontal region channels (9, 11, 22, 24, 33, 122, and 124), which were selected on the basis of the standard 10-20 system [42].

1) DWT (DISCRETE WAVELET TRANSFORM)

The discrete wavelet transform (DWT) is a stable and suitable method for signal transformation in signal processing applications. In past studies, Fast Fourier transform (FFT) was frequently used to extract features from EEG signals. EEG signals are stochastic in nature and have time-varying properties that make them non-stationary even over short durations. Unlike FFT, the DWT is suitable for non-stationary signals as it decomposes the signal into time and frequency simultaneously. Therefore, DWT is the better technique when one wants to decompose a signal into many frequency components [2]. DWT also acts as a band-pass filter and can be used to divide EEG signals into different frequency bands. The DWT can segregate (decimate) the data using the dilation and scaling factors from any mother wavelet function. The level of decimation depends upon the sampling frequency of the given signal. In this experiment, the decimation number was set to the fifth level.

2) FEATURE EXTRACTION USING DWT

The multimedia animations that we used were categorized as class two, meaning that they were associated with learning states related to working memory. Because working memory resides in the frontal and parietal region of the brain, we only analyzed the data from frontal and parietal channels.

A mathematical description for channel selection is given below.

EEG channel Function $= x[n]$

$$x[n] = \left\{ \begin{array}{l} x_{Fp1}[n], x_{Fp2}[n], x_{F3}[n], x_{F4}[n], x_{Fz}[n], \\ x_{F7}[n], x_{F8}[n], x_{P3}[n], x_{Pz}[n], x_{P4}[n] \end{array} \right\} \quad (1)$$

After channel selection, DWT was applied to the EEG signals for spectral feature extraction. The following three entropy-based features were extracted: spectral entropy, approximate entropy, and sample entropy.

3) SPECTRAL ENTROPY

Spectral entropy is a measure of EEG signal complexity and can be used to detect the strength of the cognitive activity in the signal. The EEG spectral entropy ($S_E(F)$) [43] was calculated for baseline and learning states using Equation (2), which was then normalized to the range 0–1.

$$S_E(F) = -\frac{1}{\log N_u} \sum_u P_u(F) \log_e P_u(F) \quad (2)$$

$P_u(F)$ is the probability density function, which was estimated using the EEG signal F. The PSD is calculated with respect to total power and has a frequency range between 0.5–48 Hz for each epoch. N_u represents the total number of frequencies. $P_h(F)$ is Shannon's channel entropy and an estimation of the Shannon entropy ($S_{SH}(F)$) is given in Equation (3). Applying a histogram estimate of the probability density function,

$$S_{SH}(F) = -\sum_u P_h(F) \log_e P_h(F) \quad (3)$$

Spectral entropy is especially important for analyzing biological time series, which comprise complex dynamics. It has a great prominence in the area of non-linear analysis [44]. The numerous features of the approaches have been suggested to be useful for detecting hidden yet significant dynamic properties of physiological phenomenon.

4) APPROXIMATE ENTROPY

Approximate entropy (ApEn) was proposed as an estimator that quantifies the regularity of a time series signal, and has been successfully applied to relatively short and noisy data [45]. A large value of ApEn reflects a high degree of complexity in the time series. ApEn measures the difference in the time series between the probabilities of templates of length m, which are close to each other, and templates of length m+1, which are close to each other. Hence, ApEn indicates the probability of creating a new pattern while the dimension of templates increases. The larger the probability is, the more complex the time series. The details of the algorithm are presented in [46].

Given N data points from a time series $\{x(n)\} = x(1), x(2), \ldots, x(N)$, the level of ApEn can be calculated by the following steps:

Step 1: In t-dimensional vectors, each vector is viewed as a template. The number of templates in a vector is shown in Equation (4):

$$x_t(j) = [x(j), x(j+1), \ldots, x(j+t-1)], \quad j = 1, 2, \ldots, N-t+1 \tag{4}$$

Step 2: The distance between each template and the other templates (including itself), is denoted as $d[x_t(j), x_t(k)]$, and is computed as the maximum absolute difference between their scalar components:

$$d[x_t(j), x_t(k)] = \max_{p=0,\ldots,t-1} (|X(j+p) - x(k+p)|) \tag{5}$$

Step 3: For a given template $x_t(j)$, the number of template matches are counted, and denoted as ϕ_i^t such that the number of $j(1 \leq j \leq N-t+1)$ satisfies the distance. Then $\phi_i^t(w)$, the probability of any template $X_t(j)$ is given as:

$$\phi_i^t(w) = \frac{1}{N-t+1}\phi_i \tag{6}$$

Step 4: Then the probability of the time series $\phi^t(w)$ is calculated as:

$$\phi^t(w) = \frac{1}{N-t+1}\sum_{i=1}^{N-t+1} \ln \phi_i^t(w) \tag{7}$$

Step 5: Increase the dimension to $t+1$ and follow the steps 1-4 to compute $\phi^{t+1}(w)$

$$\phi^{t+1}(w) = \frac{1}{N-t}\sum_{i=1}^{N-t} \ln \phi_i^{t+1}(w) \tag{8}$$

Step 6: Then the approximate entropy is calculated as:

$$ApEn\,(t, w) = \lim_{N\to\infty} [\phi^t(w) - \phi^{t+1}(w)] \tag{9}$$

Setting the parameters should insure a high discrimination between baseline and learning states. In this paper t is set to be 2 and w is set to 20 percent of the standard deviation of the amplitude of the time series N, which is set to 1024.

For a finite length of data point N, the approximate entropy is estimated as

$$ApEn(t, w, N) = \phi^t(w) - \phi^{t+1}(w) \tag{10}$$

5) SAMPLE ENTROPY

The ApEn usually determines the number of counts for matching sequences and does not include instances in which ln(0) is matched. This is one bias of the of ApEn method [47]. The sample entropy (SampEn) is a technique proposed to overcome this drawback by reducing the bias. The SampEn does not include the count of self-matching. A second advantage is that it does not provide template-wise matching when using conditional probabilities. It uses the template that requires the match length m+1 and it computes the probability associated with the whole time series signals [46].

The mathematical expression for Sample entropy (SampEn) can be calculated using Eq (11) and is used to reduce the self-matching bias. All steps are the same as those for measuring the ApEn. Note that the self-matched templates are not computed in Step 3.

$$\phi^t(w) = \frac{1}{N-t+1}\sum_{i=1}^{N-t+1} \phi_i^t(w) \tag{11}$$

$$SampEn(t, w) = \lim_{N\to\infty} [\ln(\phi^t(w) - \phi^{t+1}(w))] \tag{12}$$

$$SampEn(t, w, N) = \ln[\phi^t(w) - \phi^{t+1}(w)] \tag{13}$$

In this paper t is set to 2 and w is set to 20 percent of the standard deviation of the amplitude of the time series (the same settings as those of ApEn).

6) CLASSIFICATION METHOD

Classification methods are used to classify features that are extracted using different methods. There are many popular algorithms such as LDA (Linear Discriminative Algorithm), ANN (Artificial Neural Network), and SVM (Support Vector Machine) [48] (Furdea *et al.*, 2012) [49], which were developed for implementation in certain signal processing applications. Fig. 1 shows the proposed feature extraction and classification method using baseline and learning data. Compared with the others, SVM is the most powerful and well-studied algorithm, especially for nonlinear problems, as it reduces high dimensional data to lower dimensions. SVM has provided better results in brain-modality dynamic applications and achieved good results in cognitive and mental task applications. It is based on statistical methodology and has been used to map input data into high dimensional feature-vector space, in which spatial data may be linearly separable.

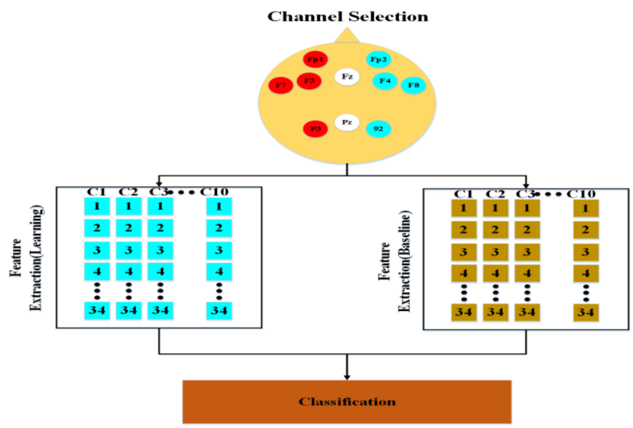
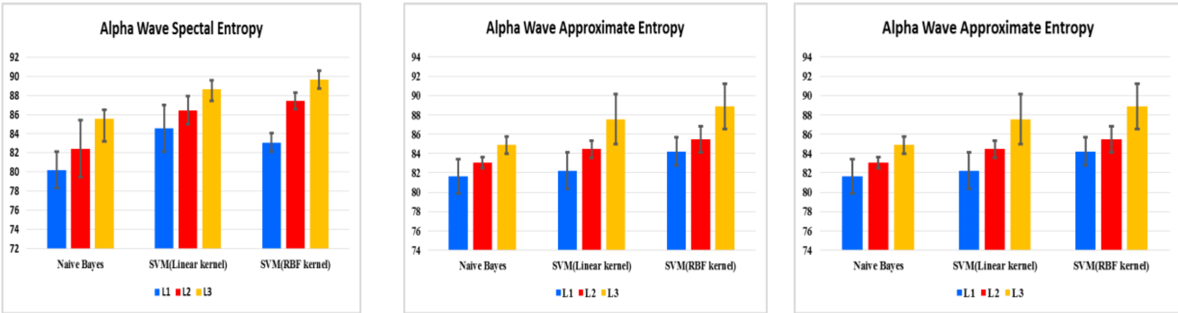


FIGURE 1. **The proposed method for eature extraction and classification of learning and baseline mental states. (C: Channel).**

7) CONNECTIVITY ANALYSIS USING PARTIAL DIRECTED COHERENCE

Connectivity is used to find interactions between two channels or two states to measure the strength between them. For example, consider two signals (or stochastic processes) X and Y with discrete time observations $x(t)$ and $y(t)$



(a) (b) (c)

FIGURE 2. Classification based on alpha-wave data. Classification accuracies for the L3 state were greater than those for the L1 and L2 states.

$t = 1, 2, \ldots\ldots, N$. Jointly, the interactions of these two signals can be described with bivariate autoregressive (ARX) models.

$$x(t) = \sum_{k=1}^{q} a_{11,k}x(t-k) + \sum_{k=1}^{q} a_{12,k}y(t-k) + e_x(t) \quad (14)$$

$$y(t) = \sum_{k=1}^{q} a_{21,k}x(t-k) + \sum_{k=1}^{q} a_{22,k}y(t-k) + e_y(t) \quad (15)$$

The linear ARX models (14) and (15) can be rewritten in matrix form and mapped to the frequency domain by Fourier transformation.

$$\begin{pmatrix} A_{11}(f) & A_{12}(f) \\ A_{21}(f) & A_{22}(f) \end{pmatrix} \begin{pmatrix} X(f) \\ Y(f) \end{pmatrix} = \begin{pmatrix} E_x(f) \\ E_y(f) \end{pmatrix} \quad (16)$$

$$\pi_{X\to Y}(f) = \sum_{k=1}^{q} a_{21,k}x(t-k)\frac{A_{21}(f)}{\sqrt{|A_{11}(f)|^2 + |A_{21}(f)|^2}} \quad (17)$$

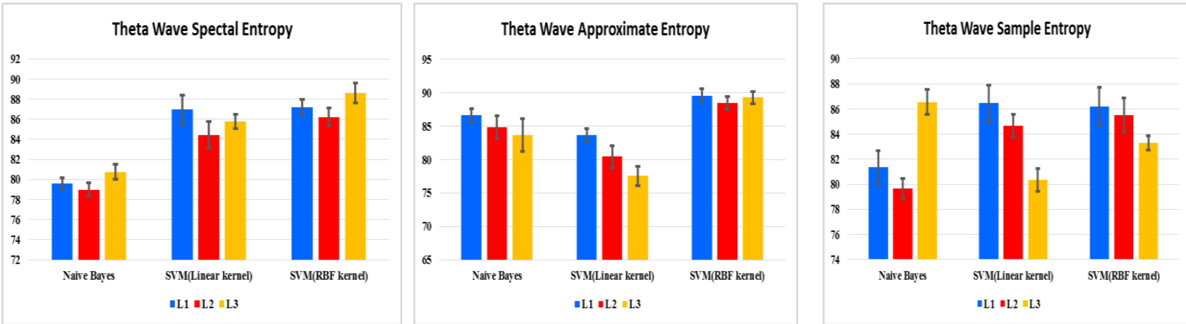
$\pi_{X\to Y}$ describes the relative coupling strength of the interaction from a signal source such as X to some signal such as Y, as compared (or normalized) to all the connections of the source to other signals. For a bivariate system, the directed interaction is described by A_{21}, and it is normalized by all the X-related terms in the ARX models (such as A_{11} and A_{21}). The magnitude of PDC then lies between zero and one.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Raw EEG is generally contaminated by artifacts such as eye blinking, muscle movement, heart rate variation, and impedance fluctuations. Brain Electrical Source Analysis (BESA) (6) software was used to remove any artifacts, and provided a clean signal for the specific tasks. After preprocessing, three entropy based features (spectral, approximate, and sample) were extracted for the baseline (EC) and learning states (L1, L2, L3). A statistical p-test was applied on the designed feature extraction technique using MANOVA software, and yielded a $p < 0.05$ for the comparison between EC and. L1-based features. Similar results were obtained when comparing the L2 and L3 states against baseline. P-tests were also used to compare the learning states with each other (L1 vs. L2, L2 vs. L3, and L3 vs. L1). The p-values show that significant differences existed between the features of the EC and learning states, indicating that they can be used to discriminate learning states from each other, and the accuracy of machine-learning algorithms. Linear kernel, RBF kernel, and Naive Bayes classifiers were applied to classify the extracted features and determine which frequencies of the EEG signal were best for discriminating between different mental states. The classisification was performed for each learing state and baseline. In the second approach, coherence values were extracted from the clean EEG signals of the learning states to assess the cognitive load and the ability to discriminate. A connectivity density factor was calculated for all learning states using all brain waves.

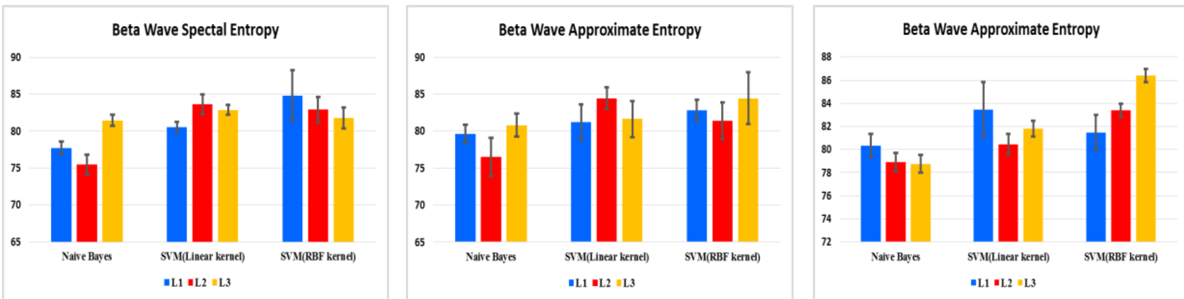
Fig. 2 shows the classification of the entropy features using alpha brain waves, based on the three classifiers for the three learning states L1, L2, and L3. The results show that for each classifier and for each entropy feature, accuracy was greatest in the L3 state, followed by the L2 and L1 states. Alpha waves have been shown to synchronize at rest and during less alert states, but become desynchronized in mental states characterized by greater alertness and higher mental effort [50]. This can explain why discrimination accuracies are higher between the L3 state and baseline than between the L2 and L1 states and baseline, where alpha waves were much synchronized. Results obtained using the combination SVM with the RBF kernel classifier produced the highest level of classification accuracy of alpha brain waves.

Fig. 3 shows the classification results for theta waves. For spectral entropy, classification accuracy did not increase sequentially for any classifier. For approximate entropy, classification accuracy increased with learning state for the Naïve-bias kernel and SVM (linear kernel). Unlike alpha waves, theta waves became less accurate in each successive learning state. This finding agrees with past research showing that theta waves are desynchronized at rest and during less alert states, but become synchronized in mental states characterized by greater alertness and higher mental effort [50]. During L1, participants were more alert and focused so theta waves synchronized more than during the L3 state in which participants required less effort to understand the learning content. Compared with the baseline state in which theta waves are desynchronized, the classifiers were thus more accurate at discriminating the L1 state than the L3 state.



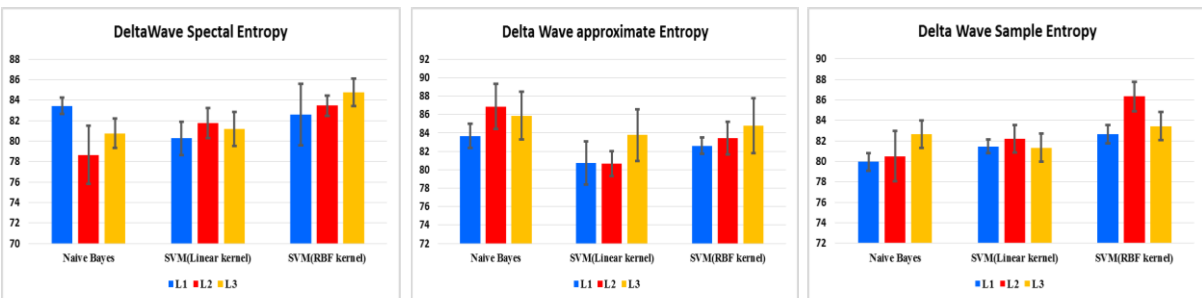
(a) (b) (c)

FIGURE 3. **Classification based on theta-wave data. The classification values do not show that the L3 state provides more accuracy than the L1 and L2 states except some cases.**



(a) (b) (c)

FIGURE 4. **Classification based on beta-wave data. The classification values do not show that the L3 state provides more accuracy than L1 and L2 states except some cases.**



(a) (b) (c)

FIGURE 5. **Classification based on delta-wave data. The classification values do not show that the L3 state provides more accuracy than the L1 and L2 states except some cases.**

Similarly, Fig. 4 shows the accuracies for beta waves. Beta waves did not yield classification accuracies that increased with learning state. Fig. 5 shows that for delta waves, no significant relationship was found between entropy features or classifiers with classification accuracy.

These overall results identify alpha waves as the best measure of cognitive load for learning-based cognitive tasks, but to some extent, theta waves also proved to be a good measure. The proposed method efficiently classifies the learning states using the extracted features. The best classification accuracies were obtained using alpha waves. Additionally, accuracy increased with learning such that the third repetition of learning material resulted in the best accuracy. Thus, the sequential accuracies differentiate the mental states during L1, L2, and L3. In the second approach, The Partial directed coherence (PDC) is calculated based on the learning-state data. PDC was calculated for the frontal and parietal channels of the 34 subjects using a sampling time of 1 s and a 250 Hz sampling frequency. The thresholding for this experiment was ≥ 0.5. L1 had PDC

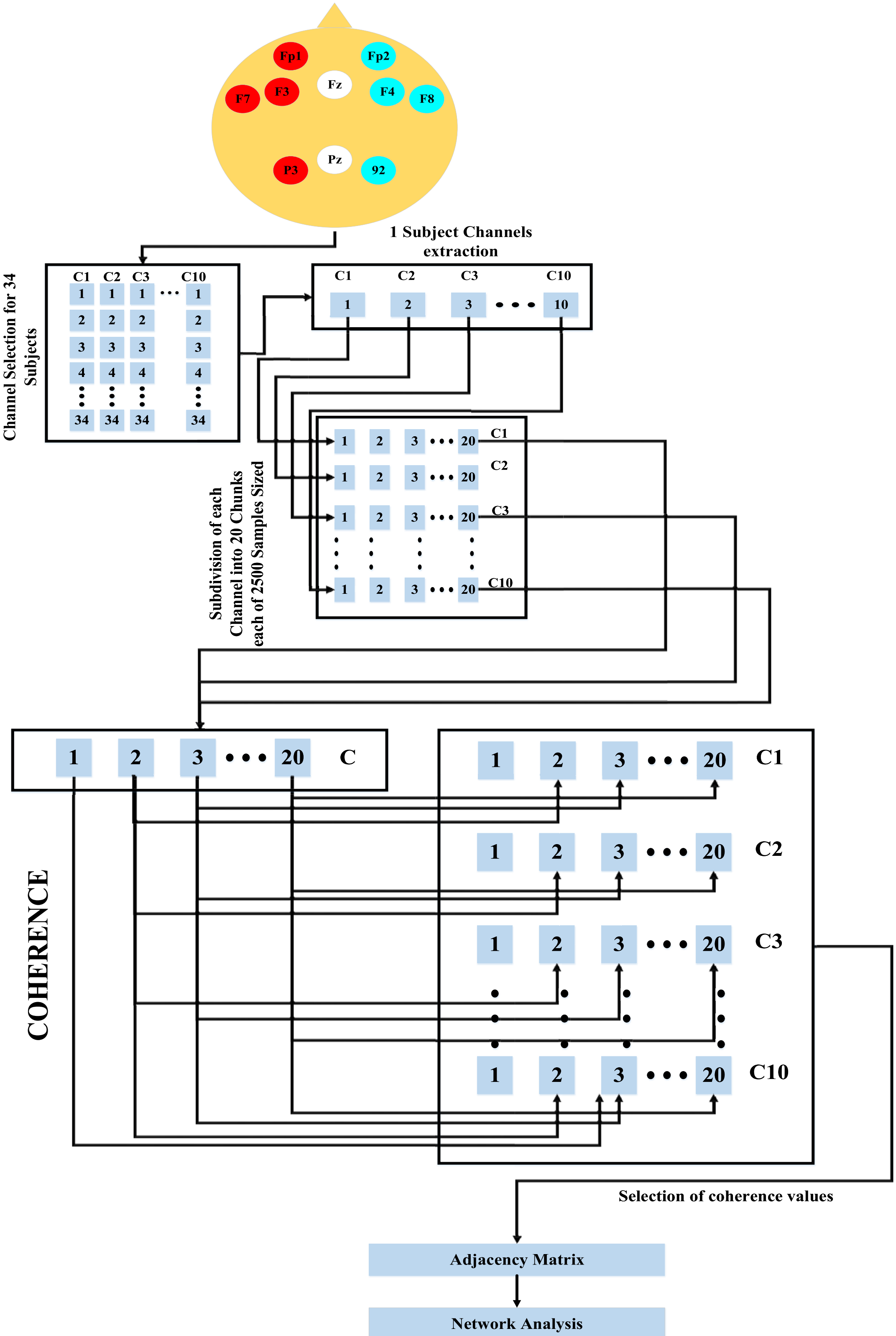


FIGURE 6. The proposed method for coherence (PDC) calculation of learning states. (C: Channel).

values ($\pi_{24\to9}$, $\pi_{11\to9}$, $\pi_{11\to22}$) for threshold 0.5 or greater. The L2 has PDC values ($\pi_{11\to9}$, $\pi_{11\to33}$, $\pi_{11\to24}$, $\pi_{33\to122}$, $\pi_{24\to122}$, $\pi_{11\to122}$, $\pi_{124\to122}$) for threshold 0.5 or greater. The L3 has PDC values ($\pi_{24\to22}$, $\pi_{11\to22}$, $\pi_{22\to9}$, $\pi_{24\to9}$, $\pi_{11\to9}$, $\pi_{124\to9}$, $\pi_{24\to122}$, $\pi_{11\to122}$, $\pi_{124\to122}$) for threshold 0.5 or greater.

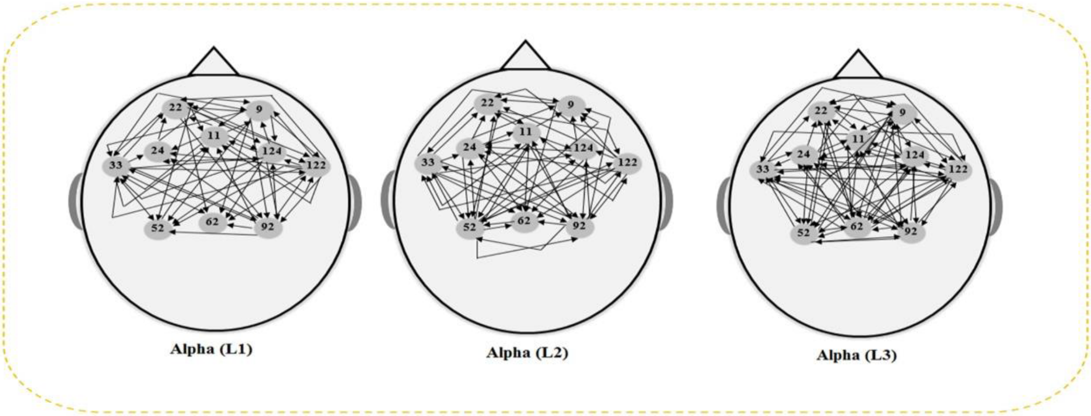


FIGURE 7. Connectivity among EEG channels during learning (L1) and rehearsal states (L2L3).

TABLE 1. Network density of each learning state based on a different number of brain waves.

Brain Waves	Alpha	Beta	Theta	Delta
Density based on L1	**114**	64	**82**	124
Density based on L2	**120**	74	58	**138**
Density based on L3	**152**	**126**	66	130

Fig. 6 shows the methodology of second method coherence to assess the cognitive load during learning task. Fig. 7 shows the flow of connectivity using partial directed coherence of the three learning states for alpha waves. The connectivity among the frontal and parietal EEG channels was calculated using PDC. Based on the PDC values for alpha waves, L3 had more connections than the other learning states. Table 1 shows the density factor of each learning state for each brain wave provided by the PDC results. For alpha waves the density factor for L3 was 150, for L2 it was 120, and for L1 it was 114. These findings imply that L3 requires less mental effort as much information has already been stored in long-term memory after the L2 and L1 stages. In PDC analysis, greater numbers of connections mean more brain areas are interconnected. Comparison with the L1 state, L2 had more connections, indicating that some of the information regarding the learning material had already been stored during L1. This is the reason why more areas of the brain are interconnected during L2 than during L1. L1 shows the fewest connections, indicating that very little information was placed in long-term memory.

V. DISCUSSION

This study assessed cognitive load measurement in multimedia learning. EEG data were collected during eye closed state and three mental states as participants viewed multimedia learning material. We extracted four frequency bands from the data using DWT (alpha, beta, theta, and delta). The spectral features (i) spectral entropy, (ii) approximate entropy and (iii) sample entropy have been extracted for the classification. Three types of classifiers (Naïve Bayes, RBF, and Linear kernel) were used to classify the cognitive states among baseline and three learning states. We analyzed the performance of each classifier for each spectral feature and each learning state.

The results showed that alpha waves had the highest classification accuracy for each of the three spectral features. These results suggest that alpha waves are the best brain waves for discriminating different learning states. The results also showed that L3 required more cognitive load than L2 or L1, while L2 had more cognitive load than L1 [51]. These findings also revealed that mental states during multimedia learning can be classified efficiently using machine learning techniques (classifiers). Hence the proposed method can be used to analyze the mental states during learning for any cognitive based task. This scheme can also be used to optimize the level of mental effort during learning-based tasks so as to better understand concepts and learning. The second part of the analysis focused on determining the effective connectivity for each of the three learning states using PDC. We found more connections during L3 than during L2 or L1, while L2 had greater number of connections than L1 for all mentioned brain waves. A greater number of connections has been shown to indicate less cognitive load [52], [53]. These findings show that by repeatedly learning the same material, participants required less cognitive load/mental effort for understanding the content. These findings indicate that this method successfully analyzed the cognitive load and different mental states during learning. Thus, both analyses show that L3 requires less cognitive load or mental effort, which indicates that proposed method is a promising way for assessing cognitive load.

VI. LIMITATIONS

In this research, two cognitive load assessment methods feature extraction and PDC are implemented. The widely used method for assessing the cognitive load in EEG is the feature extraction and classification techniques. There are limitations of this method, it gives the separate information regarding each EEG electrode. These methods are not very effective as they can only assess the cognitive load, but they do not provide the flow of information among the different brain areas during cognitive load assessment because there is no association between the feature values of the EEG electrodes. However, the human brain is functionally interconnected during any cognitive process (learning) so, a method that deals with the brain connections can provide a better overview of the information flow. Our proposed method using PDC can provide feasible solutions for cognitive load assessment. Some other limitations of the research are: it does not discriminate the different types of cognitive load like germane, intrinsic and extraneous and there is no information is presented regarding the outcomes of the learning as the stimuli based on a multimedia learning task.

VII. CONCLUSION

An EEG-based direct assessment of cognitive load in multimedia learning has been discussed in this work. The work demonstrated which EEG frequency bands, and which methods were best for assessing the cognitive load. We also validated the framework. The results showed that alpha waves can be used to estimate the cognitive load imposed by multimedia learning, and that repeated exposure to the material leads to less cognitive load. These findings indicate that entropy-based feature extraction methods and PDC are able to analyze different cognitive states during learning. Future work will explore the methods used in this work using greater numbers of channels over different brain regions to assess cognitive loads for a constant workload.

REFERENCES

[1] F. Paas, J. E. Tuovinen, H. Tabbers, and P. W. M. Van Gerven, "Cognitive load measurement as a means to advance cognitive load theory," *Edu. Psychol.*, vol. 38, no. 1, pp. 63–71, Mar. 2003.

[2] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using electroencephalography to measure cognitive load," *Edu. Psychol. Rev.*, vol. 22, no. 4, pp. 425–438, Dec. 2010.

[3] F. Paas, A. Renkl, and J. Sweller, "Cognitive load theory and instructional design: Recent developments," *Edu. Psychol.*, vol. 38, no. 1, pp. 1–4, Mar. 2003.

[4] A. Baddeley, "Working memory: The interface between memory and cognition," *J. Cognit. Neurosci.*, vol. 4, no. 3, pp. 281–288, Jul. 1992.

[5] G. S. Halford, R. Baker, J. E. McCredden, and J. D. Bain, "How many variables can humans process?" *Psychol. Sci.*, vol. 16, no. 1, pp. 70–76, Jan. 2005.

[6] J. Sweller, "Cognitive load theory, learning difficulty, and instructional design," *Learn. Instruct.*, vol. 4, no. 4, pp. 295–312, Jan. 1994.

[7] N. Kumar and J. Kumar, "Measurement of cognitive load in HCI systems using EEG power spectrum: An experimental study," *Proc. Comput. Sci.*, vol. 84, pp. 70–78, Dec. 2016.

[8] F. G. W. C. Paas and J. J. G. Van Merriënboer, "Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach," *J. Edu. Psychol.*, vol. 86, no. 1, pp. 122–133, 1994.

[9] B. Yin, F. Chen, N. Ruiz, and E. Ambikairajah, "Speech-based cognitive load monitoring system," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2008, pp. 2041–2044.

[10] B. Yin, N. Ruiz, F. Chen, and M. A. Khawaja, "Automatic cognitive load detection from speech features," in *Proc. 19th Austral. Conf. Comput.-Human Interact., Entertaining User Interfaces*, 2007, pp. 249–255.

[11] Y. Shi, N. Ruiz, R. Taib, E. Choi, and F. Chen, "Galvanic skin response (GSR) as an index of cognitive load," in *Proc. Extended Abstracts Human Factors Comput. Syst. (CHI)*, 2007, pp. 2651–2656.

[12] R. W. Backs and L. C. Walrath, "Eye movement and pupillary response indices of mental workload during visual search of symbolic displays," *Appl. Ergonom.*, vol. 23, no. 4, pp. 243–254, Aug. 1992.

[13] N. V. Thakor and S. Tong, "Advances in quantitative electroencephalogram analysis methods," *Annu. Rev. Biomed. Eng.*, vol. 6, no. 1, pp. 453–495, Aug. 2004.

[14] A. Murata, "An attempt to evaluate mental workload using wavelet transform of EEG," *Human Factors, J. Human Factors Ergonom. Soc.*, vol. 47, no. 3, pp. 498–508, Oct. 2005.

[15] U. Orhan, M. Hekim, and M. Ozer, "EEG signals classification using the K-means clustering and a multilayer perceptron neural network model," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13475–13481, Sep. 2011.

[16] N.-Y. Liang, P. Saratchandran, G.-B. Huang, and N. Sundararajan, "Classification of mental tasks from eeg signals using extreme learning machine," *Int. J. Neural Syst.*, vol. 16, no. 1, pp. 29–38, Feb. 2006.

[17] O. A. Rosso *et al.*, "Wavelet entropy: A new tool for analysis of short duration brain electrical signals," *J. Neurosci. Methods*, vol. 105, no. 1, pp. 65–75, Jan. 2001.

[18] H. Garry, B. McGinley, E. Jones, and M. Glavin, "An evaluation of the effects of wavelet coefficient quantisation in transform based EEG compression," *Comput. Biol. Med.*, vol. 43, no. 6, pp. 661–669, Jul. 2013.

[19] T. van Gog, F. Paas, and J. J. G. van Merriënboer, "Effects of process-oriented worked examples on troubleshooting transfer performance," *Learn. Instruct.*, vol. 16, no. 2, pp. 154–164, Apr. 2006.

[20] J. Sweller, "Element interactivity and intrinsic, extraneous, and germane cognitive load," *Edu. Psychol. Rev.*, vol. 22, no. 2, pp. 123–138, Jun. 2010.

[21] G. Monaci, P. Vandergheynst, and F. T. Sommer, "Learning bimodal structure in audio—Visual data," *IEEE Trans. Neural Netw.*, vol. 20, no. 12, pp. 1898–1910, Dec. 2009.

[22] L. Guo, D. Rivero, J. Dorado, J. R. Rabuñal, and A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks," *J. Neurosci. Methods*, vol. 191, no. 1, pp. 101–109, Aug. 2010.

[23] U. R. Acharya, S. V. Sree, P. C. A. Ang, R. Yanti, and J. S. Suri, "Application of non-linear and wavelet based features for the automated identification of epileptic eeg signals," *Int. J. Neural Syst.*, vol. 22, no. 2, p. 1250002, Apr. 2012.

[24] T. Gandhi, B. K. Panigrahi, and S. Anand, "A comparative study of wavelet families for EEG signal classification," *Neurocomputing*, vol. 74, no. 17, pp. 3051–3057, Oct. 2011.

[25] M. Taghizadeh-Sarabi, M. R. Daliri, and K. S. Niksirat, "Decoding objects of basic categories from electroencephalographic signals using wavelet transform and support vector machines," *Brain Topograph.*, vol. 28, no. 1, pp. 33–46, Jan. 2015.

[26] C. Vidaurre, N. Krämer, B. Blankertz, and A. Schlögl, "Time domain parameters as a feature for EEG-based brain–computer interfaces," *Neural Netw.*, vol. 22, no. 9, pp. 1313–1319, Nov. 2009.

[27] E. D. Übeyli, "Lyapunov exponents/probabilistic neural networks for analysis of EEG signals," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 985–992, Mar. 2010.

[28] R. W. Thatcher, D. North, and C. Biver, "EEG and intelligence: Relations between EEG coherence, EEG phase delay and power," *Clin. Neurophysiol.*, vol. 116, no. 9, pp. 2129–2141, Sep. 2005.

[29] M. Hariharan, V. Vijean, R. Sindhu, P. Divakar, A. Saidatul, and S. Yaacob, "Classification of mental tasks using stockwell transform," *Comput. Elect. Eng.*, vol. 40, no. 5, pp. 1741–1749, Jul. 2014.

[30] S. Noshadi, V. Abootalebi, M. T. Sadeghi, and M. S. Shahvazian, "Selection of an efficient feature space for EEG-based mental task discrimination," *Biocybern. Biomed. Eng.*, vol. 34, no. 3, pp. 159–168, 2014.

[31] S. M. Hosni, M. E. Gadallah, S. F. Bahgat, and M. S. AbdelWahab, "Classification of EEG signals using different feature extraction techniques for mental-task BCI," in *Proc. Int. Conf. Comput. Eng. Syst.*, 2007, pp. 220–226.

[32] J.-Z. Xue, H. Zhang, C.-X. Zheng, and X.-G. Yan, "Wavelet packet transform for feature extraction of EEG during mental tasks," in *Proc. Int. Conf. Mach. Learn.*, Nov. 2003, pp. 360–363.

[33] L. Zhiwei and S. Minfen, "Classification of mental task EEG signals using wavelet packet entropy and SVM," in *Proc. 8th Int. Conf. Electron. Meas. Instrum.*, 2007, pp. 3-906–3-909.

[34] C.-J. Lin and M.-H. Hsieh, "Classification of mental task from EEG data using neural networks based on particle swarm optimization," *Neurocomputing*, vol. 72, nos. 4–6, pp. 1121–1130, Jan. 2009.

[35] G. Rodríguez-Bermúdez, P. J. García-Laencina, J. Roca-González, and J. Roca-Dorda, "Efficient feature selection and linear discrimination of EEG signals," *Neurocomputing*, vol. 115, pp. 161–165, Sep. 2013.

[36] S. Karkare, G. Saha, and J. Bhattacharya, "Investigating long-range correlation properties in EEG during complex cognitive tasks," *Chaos, Solitons Fractals*, vol. 42, no. 4, pp. 2067–2073, Nov. 2009.

[37] H. U. Amin *et al.*, "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Austral. Phys. Eng. Sci. Med.*, vol. 38, no. 1, pp. 139–149, Mar. 2015.

[38] O. David, D. Cosmelli, D. Hasboun, and L. Garnero, "A multitrial analysis for revealing significant corticocortical networks in magnetoencephalography and electroencephalography," *NeuroImage*, vol. 20, no. 1, pp. 186–201, Sep. 2003.

[39] L. Astolfi *et al.*, "Assessing cortical functional connectivity by linear inverse estimation and directed transfer function: Simulations and application to real data," *Clin. Neurophysiol.*, vol. 116, no. 4, pp. 920–932, Apr. 2005.

[40] C. E. Biazoli *et al.*, "Application of partial directed coherence to the analysis of resting-state EEG-fMRI Data," *Brain Connectivity*, vol. 3, no. 6, pp. 563–568, Dec. 2013.

[41] H. U. Amin, A. S. Malik, N. Kamel, W.-T. Chooi, and M. Hussain, "P300 correlates with learning & memory abilities and fluid intelligence," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, p. 87, Dec. 2015.

[42] U. Herwig, P. Satrapi, and C. Schönfeldt-Lecuona, "Using the international 10–20 EEG system for positioning of transcranial magnetic stimulation," *Brain Topograph.*, vol. 16, no. 2, pp. 95–99, 2003.

[43] B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova, and G. B. Boylan, "A comparison of quantitative EEG features for neonatal seizure detection," *Clin. Neurophysiol.*, vol. 119, no. 6, pp. 1248–1261, Jun. 2008.

[44] L. I. Aftanas, N. V. Lotova, V. I. Koshkarov, V. L. Pokrovskaja, S. A. Popov, and V. P. Makhnev, "Non-linear analysis of emotion EEG: Calculation of kolmogorov entropy and the principal Lyapunov exponent," *Neurosci. Lett.*, vol. 226, no. 1, pp. 13–16, Apr. 1997.

[45] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Nat. Acad. Sci. USA*, vol. 88, no. 6, pp. 2297–2301, Mar. 1991.

[46] H. M. Al-Angari and A. V. Sahakian, "Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 10, pp. 1900–1904, Oct. 2007.

[47] L.-L. Chen, J. Zhang, J.-Z. Zou, C.-J. Zhao, and G.-S. Wang, "A framework on wavelet-based nonlinear features and extreme learning machine for epileptic seizure detection," *Biomed. Signal Process. Control*, vol. 10, pp. 1–10, Mar. 2014.

[48] X. Li, X. Chen, Y. Yan, W. Wei, and Z. Wang, "Classification of EEG signals using a multiple kernel learning support vector machine," *Sensors*, vol. 14, no. 7, pp. 12784–12802, Jul. 2014.

[49] D. Garrett, D. A. Peterson, C. W. Anderson, and M. H. Thaut, "Comparison of linear, nonlinear, and feature selection methods for EEG signal classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 11, no. 2, pp. 141–144, Jun. 2003.

[50] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis," *Brain Res. Rev.*, vol. 29, nos. 2–3, pp. 169–195, Apr. 1999.

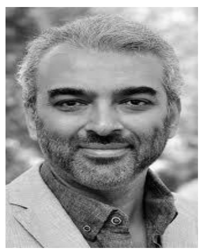
[51] P. Zarjam, J. Epps, and N. H. Lovell, "Characterizing mental load in an arithmetic task using entropy-based features," in *Proc. 11th Int. Conf. Inf. Sci., Signal Process. Their Appl. (ISSPA)*, Jul. 2012, pp. 199–204.

[52] C. H. Y. Fu *et al.*, "Modulation of effective connectivity by cognitive demand in phonological verbal fluency," *NeuroImage*, vol. 30, no. 1, pp. 266–271, Mar. 2006.

[53] D. Huang *et al.*, "Combining partial directed coherence and graph theory to analyse effective brain networks of different mental tasks," *Frontiers Human Neurosci.*, vol. 10, p. 235, May 2016.

MOONA MAZHER received the bachelor's degree in telecommunication engineering from Pakistan in 2010. She is currently pursuing the master's degree in electrical and electronics engineering from Universiti Teknologi PETRONAS, Malaysia. Her research interests include biomedical signal processing, machine learning, and optimization of deep learning algorithms.

AZRINA ABD AZIZ (M'07) received the bachelor's degree (Hons.) in electrical and electronic engineering from the University of Queensland, Australia, in 1997, the M.Sc. degree in system level integration from the Institute for System Level Integration, U.K., in 2003, and the Ph.D. degree in electrical and computer systems engineering from Monash University, Melbourne, Australia, in 2013. She is currently a Lecturer with the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia, where she is also a member of the Centre for Intelligent Signal and Imaging Research Group. Her recent research interests focus on energy-efficient techniques for topology control in wireless sensor networks and wireless body area networks for biomedical applications and medical imaging. She is a member of the Board of Engineers Malaysia.



AAMIR SAEED MALIK received the Ph.D. degree from the Gwangju Institute of Science and Technology, South Korea. He is currently serving as an Associate Professor with Universiti Teknologi PETRONAS, Malaysia. He has authored three books, over 50 papers in reputed journals, and over 100 papers in conferences. His research interest includes biomedical engineering, computer vision, pattern recognition, image processing, depth estimation, and visual surveillance. He has received awards in various international exhibitions. He served on technical committees of various conferences and publications. He is the Editor-in-Chief of the *RESINNEX*, a UTP research publication.

HAFEEZ ULLAH AMIN received the B.S. degree (Hons.) in information technology and the M.S. degree in computer science from Pakistan in 2006 and 2009, respectively, and the Ph.D. degree in 2017. Since 2011, he has been involved in research on memory retention with Universiti Teknologi PETRONAS, Malaysia, where he is currently a Research Scientist with the Center for Intelligent Signal and Imaging Research. He has published his work in peer-reviewed journals and indexed conferences. His research interests include EEG/ERP signals analysis, long-term memory assessment, and intelligent information system. He is a member of the IEEE Engineering in Medicine and Biology Society, and the Society for Neuroscience.