Expert Systems With Applications 112 (2018) 138–147

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Brain-computer interface for workload estimation: Assessment of mental efforts in learning processes

Amin Zammouri[a,*], Abdelaziz Ait Moussa[b], Yassine Mebrouk[c]

[a] Euro-Mediterranean University of Fez, Fez 30070, Morocco
[b] Department of Computer Science, Faculty of Sciences, Mohammed First University, Oujda 60000, Morocco
[c] Service de Neurologie, Centre Hospitalier Universitaire Mohammed VI, Oujda 60000, Morocco

ARTICLE INFO

Article history:
Received 13 July 2017
Revised 8 May 2018
Accepted 10 June 2018
Available online 18 June 2018

Keywords:
Workload
Cognitive task
EEG
BCI
Learning process

ABSTRACT

To assess the current mental state of an individual, several monitoring systems have been developed. In this paper, we explore the possibility to exploit information recorded noninvasively from the human cortex to develop a brain-computer interface (BCI) able to estimate brain workload and the mental efforts during a cognitive task. The EEG-based workload classifier presented in this paper combines a power spectral density (PSD) analysis and a statistical criterion. The proposed classifier is applied in the context of distance education and online course platform through two experimental protocols. The first one proposes solving a set of matrices products using pen and paper, while the second one proposes answering problems of logic mathematics on a computer-based learning environment. Experimental results show that the averaged accuracy of distinguishing changes in the theta [4–7 Hz] (θ) band is 79%. For the alpha band [8–11 Hz] (α) the averaged accuracy reached 78%. Based on this classifier, we demonstrate that θ and α powers in central, and posterior sites decrease with the increase in difficulty level of the cognitive task. The accuracy of correct decisions obtained from our results are significantly enhanced while comparing our investigation to some similar works from literature. In the realm of intelligent expert systems, our work results represent a first step to an implementation of an intelligent system for the evaluation of reeducative therapies to be used in physiotherapy centers for children with cognitive disorders as it is the case for Cerebral Palsy.

1. Introduction

Learner's attention and mental effort are the primary factors for acquiring knowledge in an educative context. Hence, estimating mental efforts of learners in real time can help designing adaptive learning techniques, especially for disabled people (Anderson, Betts, Ferris, & Fincham, 2012). Previous researches showed that various physiological factors co-vary with mental effort levels (Healey, 2005). Among these factors, we distinguish Heart Rate Variation (HRV), Galvanic Skin Response (GSR) and the Electro-encephalogram (EEG) activity (Wobrock et al., 2015). In recent years, more and more researches have been focused on the development of automatic systems to assess mental effort and the quantity of engaged cognitive sources during a cognitive task (Huang et al., 2016; Roy, 2016; Spüler et al., 2016; Zammouri, 2016). These works are based on the physiological information associated with brain activity, here in EEG. Relying on EEG is justified by the visible effects induced by changes in cognitive state. The increase and decrease in the level of cognitive load is the result of many complex factors interacting together. The user's workload can be defined either as the imposed load of working memory or the number of works to perform concurrently. More generally it is defined as the measurement of the amount of mental resources involved in a cognitive task. The use of the EEG measurements to assess the mental workload takes advantage of the spectral aspect of the electrical brain signal. According to Başar et al. (2001), the electrical brain activity generates different rhythms. These rhythms represent a continuum of waves: Delta [0.5–3 Hz] (δ), Theta [4–7 Hz] (θ), Alpha [8–11 Hz] (α), Beta [12–30 Hz] (β) and Gamma [> 30 Hz] (γ). For Klimesch (1999) and Andreassi (1995) θ and α oscillations are sensitive to task difficulties. According to Holm, Lukander, Korpela, Sallinen, and Müller (2009) the increase of activity demands increases the θ oscillations in the frontal brain area (electrode Fz) and decreases the α oscillations in the parietal brain area (electrode Pz). Apart from brain rhythms, some works used Event Related Potentials (ERPs) as indicators of mental states. These potentials represent specific responses to cognitive events called stimuli. Amongst the widely known evoked potentials, P300

* Corresponding author.
E-mail addresses: a.zammouri@ueuromed.org (A. Zammouri), a_aitmoussa@yahoo.fr (A. Ait Moussa), mebrouky@gmail.com (Y. Mebrouk).

https://doi.org/10.1016/j.eswa.2018.06.027
0957-4174/

evoked component has justified its robustness to assess mental states and workload through its sensitivity to the brain processing competencies (Holm et al., 2009; Röy, Charbonnier, Campagne, & Bonnet, 2016; van Dinteren, Arns, Jongsma, & Kessels, 2014).

Through works from the literature, the development of a brain-computer interface (BCI) provides a system with possibility of communication based only on brain signals. Hence, in the context of assistance and rehabilitation, BCI systems can be a reliable tool to assist, improve or repair some of the human cognitive functions (Bell, Shenoy, Chalodhorn, & Rao, 2008; Pires, Nunes, & Castelo-Branco, 2012; Wang, Jin, Zhang, & Wang, 2014; Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002; Yasui, 2009). Amongst the recent innovative tools designed in the realm of estimating mental states and workload, passive Brain-Computer Interfaces (pBCI) exploit spontaneous brain activity for implicit brain-machine interactions. While most researches in BCI are designed for direct control with disabled people, pBCI use brain measurements as an additional input in order to reflect information about the user's cognitive state (Farwell, 1988; Leeb et al., 2013). This term, passive BCI, was introduced for the first time by Cutrell and Tan (2008) and refers to an implicit interaction with a computer based on brain activity. The term implicit brain-computer interface reveals the relationship that exists between implicit interactions in general and interactions based on brain activity. Current brain activity measurements allow accessing to different types of implicit informations on the user's mental state such as the cognitive workload.

In this work, and in relation to the pBCI context, we aim at developing a BCI to assess and estimate learners' mental states and workload based on EEG signals. The approach presented in this paper exploits two experimental protocols. The first protocol aims at measuring the EEG signals from four subjects while solving matrices products using pen and paper. In the second protocol eight subjects answered multi-choices questions of mathematical logic using a computer-based learning environment. In this investigation we try to address the issue of estimating workload changes and to highlight brain regions that interact during a learning process. The proposed solution is based on a combination of power spectral density analysis and the Student's law applied to θ and α brain rhythms. Through our designed classifier we demonstrate how EEG data could be used for distinguishing different levels of an individual's workload especially in the context of a learning process. The developed classifier was tested using EEG data from the first protocol and validated on the data set from the second one. The major faced difficulty in estimating and classifying levels of the cognitive workload was related to the fact that there are many states and not a few wholly states.

In relation to the expert systems, this work results, in terms of an individual's mental effort levels classification, allow to introduce a new aspect of intelligence which improves the human operator's activities in general. Indeed, in the context of road safety, the proposed classifier can be used to identify and detect the departure in a state of falling asleep or drowsiness. Unlike the currently existing systems (such those based on movements detection in the car or those based on eyes closure detection through a camera), the use of the proposed classifier makes it possible to detect pre-sleep mental states. Such an approach will make it possible to design alert systems or remediation systems before going in a sleep state.

On the intersection of biotechnologies, health care and Intelligent Tutoring Systems (ITS), the proposed classifier, allows to design new ITS whose expert module intelligence derives its robustness from the learner's current mental effort consideration. Such an intelligence enables the ITS expert module to accurately select and adapt the learning content for the learner based on his current mental state. Such an expert and intelligent system is of a crucial importance and of an immediate demand in the context of physiotherapy and rehabilitation especially when it concerns a disease of cognitive disorders as in the case of Cerebral Palsy (CP). Such an ITS enables the concerned specialists to create adapted learning scenarios in order to properly evaluate their reeducational therapies according to patients.

The paper is organized as follows. Section 2 provides a background of works in relation to mental states and workload recognition. In Sections 3 and 4, we present the adopted model for estimating the workload. Experimental results are presented and discussed in Sections 5 and 6. Finally conclusions are drawn in the last section.

2. Background and related works

The brain workload is considered as the memory work imposed load during a cognitive task (Hart, 1988). Analyzing this brain workload using the EEG signals was the subject of many research works in psychology. The theory of cognitive load was developed for the first time in 1980 by the psychologist John Sweller (Sweller, 1988). Through two types of cognitive load, Sweller aimed to explain success and failure in learning and cognitive tasks. The first cognitive load type, known as intrinsic cognitive load, describes the characteristics of the learning content interactivity level. The extrinsic cognitive load, in turn, corresponds to the way of presenting the learning content. Sweller compared this load to anything that may disrupt the learning process. Gevins (2003) introduced an approach in performing n-back tests. Gevins aimed to extract spectral characteristics of θ and α rhythms in a four-seconds time-window. Extracted information are then processed with a neural-network. Using this approach, Gevins has discriminated three various levels of cognitive workload. In the n-back cognitive task proposed by Gevins, each user saw a sequence of letters on screen. These letters are displayed one by one every two seconds. For each letter the user had to determine if the displayed letter was already displayed in the N previous letters or was a new displayed one.

Holm et al. (2009) has determined the $\theta(Fz)/\alpha(Pz)$ ratio as a sensitive indicator of the overall brain load. According to Holm, the α power could differentiate task demand levels from each other. However, Holm has considered that the α power is not an optimal method for estimating the overall workload, since according to her results, it increases when an alert subject is working at easy task demand level, as well as when a subject is engaged in a complex multitask. In contrast to the findings of Postma, Schellekens, Hanson, and Hoogeboom (2005), who was the first to define the $\theta(Fz)/\alpha(Pz)$ ratio in load estimation, Holm has showed that internal state of the subject has a strong effect on the ratio. Indeed, time awake and sleep restriction significantly increase the $\theta(Fz)/\alpha(Pz)$ ratio. In addition to that, a comparison of the P300 amplitude and $\theta(Fz)/\alpha(Pz)$ ratio made by Holm, has showed that the ratio performed well in differentiating the task demand levels. In this context of designing an index for describing changes in the cognitive load, our recent research work (Zammouri et al., 2017) is interested in studying this index established by Postma an Holm. Studying this index is interested in its behavior when applying it to EEG data measured in a context of performing cognitive exercises of different difficulties. Such a comparison revealed that the Postma and Holm ratio does not keep the same behavior when varying the cognitive task difficulty. In such a situation, results presented in Zammouri et al. (2017) describe the $z = \theta(P4)/\alpha(F4)$ index as an improvement of the Postma and Holm index. Applied on two test sessions, the z index keeps the same behavior and decreases when moving from the first session to the second one, i.e., when increasing the cognitive task difficulty. Results from (Zammouri et al., 2017) show that when increasing the cognitive task difficulty, this

	First matrix	Second matrix
Product 1	$M_1 = (20)$	$M_2 = (7)$
Product 2	$M_1 = (4 \quad 2)$	$M_2 = \begin{pmatrix} -2 \\ 2 \end{pmatrix}$
⋮	⋮	⋮
Product N	$M_1 = \begin{pmatrix} -5 & 16 & 2 & 19 & 14 & 40 \\ 10 & -1 & -8 & 3 & 11 & 12 \\ 15 & 0 & 15 & 0 & -5 & -3 \\ 25 & -50 & 50 & 28 & 22 & 11 \\ -7 & 5 & 18 & -34 & -7 & -3 \\ -32 & 0 & -20 & 6 & 14 & 7 \end{pmatrix}$	$M_2 = \begin{pmatrix} -5 & 40 & 2 & 19 & 14 & 16 \\ 10 & 12 & -8 & 3 & 11 & -1 \\ 15 & -3 & 15 & 0 & -5 & 0 \\ 25 & 11 & 50 & 28 & 22 & -50 \\ -7 & -3 & 18 & -34 & -7 & 5 \\ -32 & 7 & -20 & 6 & 14 & 0 \end{pmatrix}$

Fig. 1. Examples of the used matrices products.

phenomenon is accompanied by a decrease in the amount of the θ wave in the parietal brain area.

Putze, Jarvis, and Schultz (2010) proposed a multimodal approach for cognitive load recognition. In addition to EEG data, Putze have measured the galvanic skin response and breathing. These data were then processed by a classifier based on a one-minute moving window. This approach was evaluated on subjects in a driving simulator performing a lane change task while solving a cognitive secondary task. In this context of driving, authors in Pal et al. (2008) proposed a mental fatigue estimation approach based, at a first time, on measuring the driver's mental state of rest. This rest state is considered as the reference state. This reference state is represented by two models describing the variations of θ and α brain waves in the posterior brain area. During the user's driving experience, the proposed approach is to measure the distance away from the reference state. Such an approach is based on using large time windows to construct the appropriate input matrices for computing the Mahalanobis distance.

Apart from these approaches, recent works focused on developing and enhancing methods of workload estimation based on spatial filters. Dijksterhuis, de Waard, Brookhuis, Mulder, and de Jong (2013) assessed changes in drivers' visiomotor workload using the Common Spatial Pattern (CSP) and the Fisher's Linear Discriminant Analysis (FLDA). Dijksterhuis's classification accuracies were obtained from lower EEG ranges and reached 80%. In the same idea, Mühl, Jeunet, and Lotte (2014) used CSP filtering with 2 s analysis windows to classify mental workload in an n-back task. Based on the CSP, Mühl obtained 73% of correct classification. In this context, Röy et al. (2016) compared performances of using spectral markers and ERPs to estimate mental workload while performing a Sternberg memory task. Roy included CSP, Canonical Correlation Analysis (CCA) and FLDA. The obtained results showed low performance (60%) while using spectral markers. Conversely, the ERP-based chain reached 91% and was stable in time.

3. Materials

3.1. Experiments

Two experimental protocols are described as follows. The first one involves solving a linear algebra problem. Participants were asked to solve a set of matrices products using pen and paper. Assuming that m and n represent respectively rows and cols of the used matrices, $(m, n) \in \{1, 2,\ldots, 6\}^2$. Matrices products were presented to participants in an increasing order of difficulty level, starting from the product of matrices of sizes $(1, 1) \times (1, 1)$ until products of sizes $(6, 6) \times (6, 6)$. An example of the used matrices is presented in Fig. 1. This cognitive task was chosen for its high visual attention and vigilance demands in calculation. On another hand, through this protocol, we try to get as close as possible to real life conditions to estimate the learner's workload during learning tasks. Upon approval, this protocol was conducted in the Neurology Department at the Mohammed Sixth Hospital Center, Morocco. Each participant visited the Neurology Department two times. In the first visit, the participant had to perform two subsequent measuring sessions. The first session aims at the measurement of the subject's brain activity during a relaxed state of 15 min with closed eyes. In the second session, the participant performed the first task of solving matrices products of low difficulty level, i.e., products of matrices of sizes from $(1, 1) \times (1, 1)$ until $(3, 3) \times (3, 3)$. During the second visit, taking place at least 24 h after the first visit, the subject performs the second task i.e., products of matrices of sizes from $(4, 4) \times (4, 4)$ until $(6, 6) \times (6, 6)$.

In the second experimental protocol, we assess the participants' workloads based on a computer-based learning environment. The cognitive task aims at answering 20 multiple-choice questions covering the mathematical and the logic arithmetic. Three difficulty levels were proposed through the 20 questions which were presented to the participant randomly. In the first level, denoted by l_1, subjects were asked to complete the missing element in a numerical series. In the second level, denoted by l_2, subjects were asked to solve problems such as conversion of time and temperature units. Finally, in the third level, denoted by l_3, we presented to subjects questions which require problem solving using different arithmetic operations. Each of the 20 questions from this experimental protocol was presented separately with four answer proposals. A 5-min relaxation session, with closed eyes, was applied to each participant before starting the experiment. This experimental protocol was conducted at the Versailles Systems Engineering Laboratory, France. Applying relaxation sessions in both experimental protocols aims at reducing effects of the participant's brain activities before starting the experiment. For example, walking to the place of experimentation could influence brain rhythms associated with motor activity.

Unlike some works similar to our's, and which have been interested in classifying the levels of cognitive load, choosing mathematical exercises as cognitive tasks of test allows getting closer from a daily life situation and proving the usefulness of a BCI as an expert system outside the laboratories. For comparison, in Brouwer et al. (2012) and Röy et al. (2016) works, the study and analysis of changes in the level of cognitive load were based on the n-back test as a cognitive task. Such a test involves evaluating the cognitive load associated with the visual memory. Specifically it consists on sequentially presenting to the user a set of letters. For each letter, the user must decide if the letter has already been presented. Despite the positioning of this cognitive exercise far from the learning context, it allows designing effective brain load classifiers.

In our work, we chose to answer questions concerning matrices algebra as well as mathematical logic and arithmetic because it allowed us to induce different difficulty levels on the test cognitive task. These difficulty levels allowed generating various levels of the participant's cognitive load. Therefore, this will increase the performance when applying our designed classifier. Hence, one should not understand that the choice of the cognitive task of our work is dependent to the desired classifier. But any other cognitive task could be adopted. We specify that we were not interested in evaluating the participant's knowledge. But, we focused our efforts on studying the behavior of the brain when varying the difficulty levels of the cognitive task. In the case of such an evaluation, it is necessary to take into account the parameter of time put on the task.

In another hand, it should be noted that theoretically, a high level of difficulty introduced into a cognitive task involves a set of brain resources which enter into interaction and consequently a high level of mental effort. This represents the postulate that we are trying to demonstrate in this paper through the adopted experimental protocols and the proposed classifier.

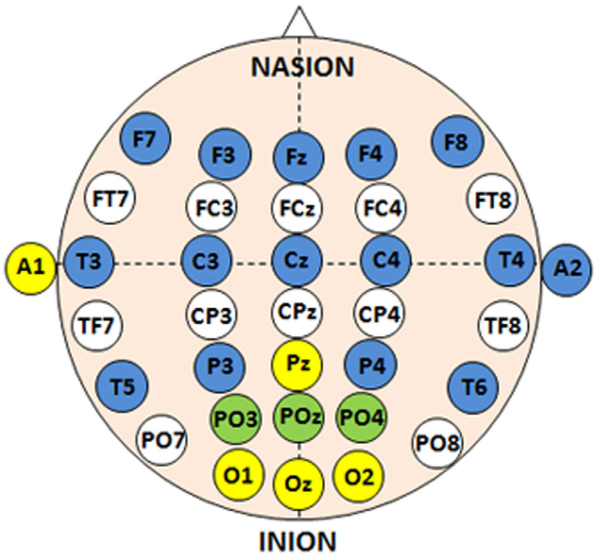

Fig. 2. Placement of used electrodes. Yellow color: electrodes used in both the two experimental setups. Blue color: electrodes used in the first experimental protocol. Green color: electrodes used in the second experimental protocol. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Participants

In the first protocol, four healthy men, right handed and aged from 23 to 33 years old, participated voluntary in the experimentations. In the second protocol, experiments included eight participants who have voluntarily accepted to participate in this study. They were all mal university students (from different countries) aged from 21 to 25 years old. In both protocols informed and anonymous consent was obtained from each participant. All participants did not present any mental or ocular troubles that may affect the experimental results.

We note that the process of selecting participants consisted of recruiting a multitude of candidates of both sexes. In order to obtain the better EEG signals clarity, we kept only candidates whose measured signals are less disturbed. These disturbances are related to the participants. The cosmetic products used by females as well as the hair gels used in both sexes reduce the conductivity of the EEG signals and disturb the measuring process. On another hand, because of the factor of stress, generated by the presence in the experiment environment, especially in females, the latter blinked frequently their eyes. This implies the presence of significant disturbances on the EEG signals. For the sake of unclear interpretation of the measured signals, we decided to keep the representative cases which we present in this paper. This selection gave a final set composed from only male participants. We also note that for reasons of ethics and availability of participants, the 21–33 age range was the most suitable for our work. One should note that the results may be different in the case of recruiting participants aged less than 10 years old for example.

3.3. Data acquisitions

EEG data set from the first protocol were recorded continuously in clinical conditions basing on the EBNeuro acquisition device with 22 electrodes placed according to the 10–20 international system as depicted in Fig. 2. The sampling frequency was set at 128 Hz. The EEG data were filtered using a pass-band filter for [1–30 Hz]. In the second protocol, the EEG data was measured continuously based on the g.Mobilab + acquisition device with eight electrodes and using a sampling rate of 256 Hz. Electrodes were placed at the occipital and parietal lobes following the extended 10–20 international system as presented in the Fig. 2. Acquisitions were performed basing on the OpenViBE (Renard et al., 2010) acquisition server.

4. Methods

4.1. Data preprocessing

Most troublesome artifacts observed in EEG are of an ocular origin. Ocular artifacts may be due to saccades, i.e., rapid eye movements (around 1000°/s), or due to eye blinks which are characterized by large amplitudes, i.e., ten times greater than EEG amplitude (Zammouri, Aitmoussa, Chevallier, & Monacelli, 2015). An effective way to deal with the ocular artifacts is by asking subjects to restrict their eye movements fixating on a stable point. However, this fixation may affect the neuroscience interpretation of the results. Therefore we must resort to other methods able to decompose the measured EEG signals. In our experiments we used the Blind Sources Separation (BSS) paradigm to overcome this problem (Jung et al., 2000; Romo Vázquez et al., 2012). For the sake of completeness, we briefly explain it.

In our work, due to the quasi-static regime we assume that the source signals arrive simultaneously on the sensors. The mixture model in this case is written as follows:

$$X = AS + N. \tag{1}$$

where $X = [X_1 X_2 \dots X_{N_e}]^T$ represents the $N_e \times N_t$ signals matrix and $S = [S_1 S_2 \dots S_{N_s}]^T$ represents the $N_s \times N_t$ original data source matrix. The sources are estimated meaning a linear transformation of signals X and that makes them as independent as possible:

$$S(t) = W^T X(t). \tag{2}$$

BSS algorithms seek the matrix W so that the $W^T A$ product is a reduced and diagonal matrix. As direct implementation of the BSS paradigm we have chosen to use Independent Component Analysis (ICA) (Zhou, 2009).

4.2. Workload classification

Brain workload is characterized by brain activities evolution in the frequency range of [1–30 Hz] (van Dinteren et al., 2014; van Gog, Kester, Nievelstein, Giesbers, & Paas, 2009). In both the two experimental protocols, a first step of artifacts rejection, using ICA, was necessary before any data processing. This filtering step aims at reducing the effects of artifacts during the computation and estimation of spectral powers in θ and α bands. We decided to focus our estimating approach on θ and α since according to many previous works these two rhythms reflect changes in the cognitive and memory performance (Klimesch, 1999; Pal et al., 2008). The introduced classifier is based on the Short Time Fourier Transform (STFT) in computing the EEG power spectrum. The STFT is calculated using the Welch (1967) periodogram method which aims at calculating an averaged periodogram with overlaps. The signal is segmented into several overlapping and equal blocks. Thus the STFT is calculated on each block. The result is the arithmetic average of the segments transform. On each EEG channel, Welch periodograms are computed on epochs lasting $d_e = 10$ s with an overlapping rate of 50%. The STFT was computed using a moving window of 2 s. This moving window duration seems appropriate in the frequency rage [1–30 Hz].

The classifier makes use of a statistical test based on the Student's law. The principal is to make a decision whether powers from two different samples are from the same class. We denote by P^* the power spectrum computed on a complete EEG signal recorded from a single channel. (*) refers to θ or α to represent their powers. The classifier consists in comparing means of powers of two epochs e_1 and e_2 of the power spectrum.

By denoting the means on epochs e_1 and e_2 by μ_{e_1} and μ_{e_2} respectively, the test of hypotheses is designed as follows: Null hypothesis (H_0) implies $\mu_{e_1} = \mu_{e_2}$ while Alternative hypothesis (H_A) corresponds to $\mu_{e_1} \neq \mu_{e_2}$.

Let P^{*,m_1} and P^{*,m_2} be the power spectrums computed from two independent EEG measures, m_1 and m_2 respectively, using the same channel. We denote by $P^{*,m_1}_{e_1}$ and $P^{*,m_2}_{e_2}$ two epochs of power spectrum collected respectively from P^{*,m_1} and P^{*,m_2}. We make the empirical conditional means as follows:

$$\overline{P^{*,m_1}_{e_1}} = \frac{1}{n_{e_1}} \sum_{i=1}^{n_{e_1}} P^{*,m_1}_{e_1}(i)$$
$$\overline{P^{*,m_2}_{e_2}} = \frac{1}{n_{e_2}} \sum_{i=1}^{n_{e_2}} P^{*,m_2}_{e_2}(i) \tag{3}$$

where n_{e_1} and n_{e_2} are respectively sizes of the tow epochs. The test involves comparing the estimated quantities $\overline{P^{*,m_1}_{e_1}}$ and $\overline{P^{*,m_2}_{e_2}}$ by taking into account variance of power values in each epoch. The computations differ according to conditional variances assumptions. In our case, values of standard deviations $\sigma(P^{*,m_1}_{e_1})$ and $\sigma(P^{*,m_2}_{e_2})$ are unknown. In order to get them we use unbiased estimators:

$$s^2\left(P^{*,m_1}_{e_1}\right) = \frac{1}{n_{e_1}-1} \sum_{i=1}^{n_{e_1}} \left(P^{*,m_1}_{e_1}(i) - \overline{P^{*,m_1}_{e_1}}\right)$$
$$s^2\left(P^{*,m_2}_{e_2}\right) = \frac{1}{n_{e_2}-1} \sum_{i=1}^{n_{e_2}} \left(P^{*,m_2}_{e_2}(i) - \overline{P^{*,m_2}_{e_2}}\right) \tag{4}$$

If we assume that the variances are identical in the two epochs a synthetic variance estimator can be designed:

$$s^2 = \frac{(n_{e_1}-1)s^2\left(P^{*,m_1}_{e_1}\right) + (n_{e_2}-1)s^2\left(P^{*,m_2}_{e_2}\right)}{n_{e_1}+n_{e_2}-2} \tag{5}$$

Under the hypothesis H_0 we define the statistic T_{obs} as follows:

$$T_{obs} = \frac{\overline{P^{*,m_1}_{e_1}} - \overline{P^{*,m_2}_{e_2}}}{\sqrt{\frac{s^2}{n_{e_1}} + \frac{s^2}{n_{e_2}}}} \tag{6}$$

The T_{obs} statistic follows a Student distribution $\mathcal{T}(\nu)$ with $\nu = n_{e_1} + n_{e_2} - 2$ degrees of freedom. For a bilateral test, the critical area is defined by:

$$|T_{obs}| \geq t_{1-\frac{\rho}{2}(\nu)} \tag{7}$$

where $t_{1-\frac{\rho}{2}(\nu)}$ is the quantile of order $1 - \frac{\rho}{2}(\nu)$ of the Student distribution. On another hand, since we are conducting N_{tests} multiple tests with a given risk ρ, an average of $\rho \times N_{tests}$ false positives is estimated. To address this issue, the classifier is complemented by the algorithm of False Discovery Rate (FDR) (Benjamini, 1995).

5. Experimental results

Here we present results with particular focus on effects of increasing the difficulty level of the cognitive task on the power spectral density of θ and α bands. In the first protocol we have chosen a set of matrices products that subjects performed through two tests. The first test was characterized by matrices of low sizes, ($(m, n) \in \{1, 2, 3\}^2$). However, in the second test we increased the sizes of matrices, i.e., $(m, n) \in \{4, 5, 6\}^2$. The presented results in Fig. 3 show that the classifier was able to well distinguish the two workload levels induced in the two tests of this protocol. In the second protocol the eight subjects were asked to answer 20 questions on logic mathematics. We have chosen these questions so that to induces three difficulty levels: l_1, l_2 and l_3. In order to better evaluating the behavior of the proposed classifier, the questions were presented to subjects in a random order of the difficulty level. The classifier was applied to each electrode of the two experimental protocols given a risk $\rho = 5\%$ to which corresponds the quantile value 2.0301 in the Student's table. Fixing the quantile value for each group of subjects implies that the workload classifying approach relies on a confidence threshold which is independent to subjects, but it is selected according to d_e duration of epochs on which the power spectrum is computed.

Table 1
Performance characteristic parameters of the developed classifier (first experimental protocol).

Subjects	θ			α		
	TPR (%)	SPC (%)	ACC (%)	TPR (%)	SPC (%)	ACC (%)
S1	100	85	90.47	66	75	71.45
S2	100	31	57.14	100	60	80.95
S3	92	50	76.19	75	100	85.71
S4	94	33	85.71	50	60	52.38
Average	96.5	49.75	77.38	72.8	73.8	72.62

Results obtained from the first experimental protocol, presented in Fig. 3, demonstrate important changes in the brain workload when moving from the first test to the second one. Changes concerning the θ wave are basically located at temporal (electrodes T3 and T6), parietal (P3 electrode) and occipital (electrodes O1 and O2) areas. For the α band, changes were mainly in brain regions: temporal (electrodes T3 and T6), central (electrodes C3, Cz and C4), parietal (electrodes P3, Pz and P4) and occipital (electrodes O1 and O2). Regarding their sensitivity to any motor activity and in order to avoid biases in the analysis, electrodes from central and frontal areas were excluded in the analysis process.

Results from the second experimental protocol are displayed in Fig. 4. One can see that, in the case of α band, values of the $|T_{obs}|$ are greater than the threshold; particularly in the occipital lobe (electrodes: Oz, POz and Iz). This reflects that changes in the brain workload took place in this brain area and which matches with the classifier outcomes from the first protocol. In turn, changes concerning the θ band were located especially at occipital brain areas (electrodes O1 and O2). Besides classifying levels of brain workload, the presented classifier allows the identification of brain areas in which there has been significant changes of workload when performing cognitive tasks used in the two experimental protocols. One should note that we take into account the assumption that when a subject started each experiment, he does not present any sign of mental fatigue.

The comparison of the workload classifier metrics using performance characteristic parameters (Hanley, 1982) revealed that the classifier performed well in detecting workload changes. The performance characteristic parameters for θ and α, calculated on all subjects, from the first experimental protocol, are presented in Table 1, while those from the second protocol are displayed in Table 2. This performance evaluation technique consists on comparing the classifier outcomes to decisions made by an expert on the same data set. The true positive rate (TPR) or detection rate is the ratio between the number of annotations of true cognitive load changes made by the classifier and the number of annotations of true changes made by the expert. The false positive rate (FPR = 1 − SPC) is the ratio of the number of false decisions made by the classifier and the number of the expert's annotations concerning non-changes of the cognitive workload. These rates are

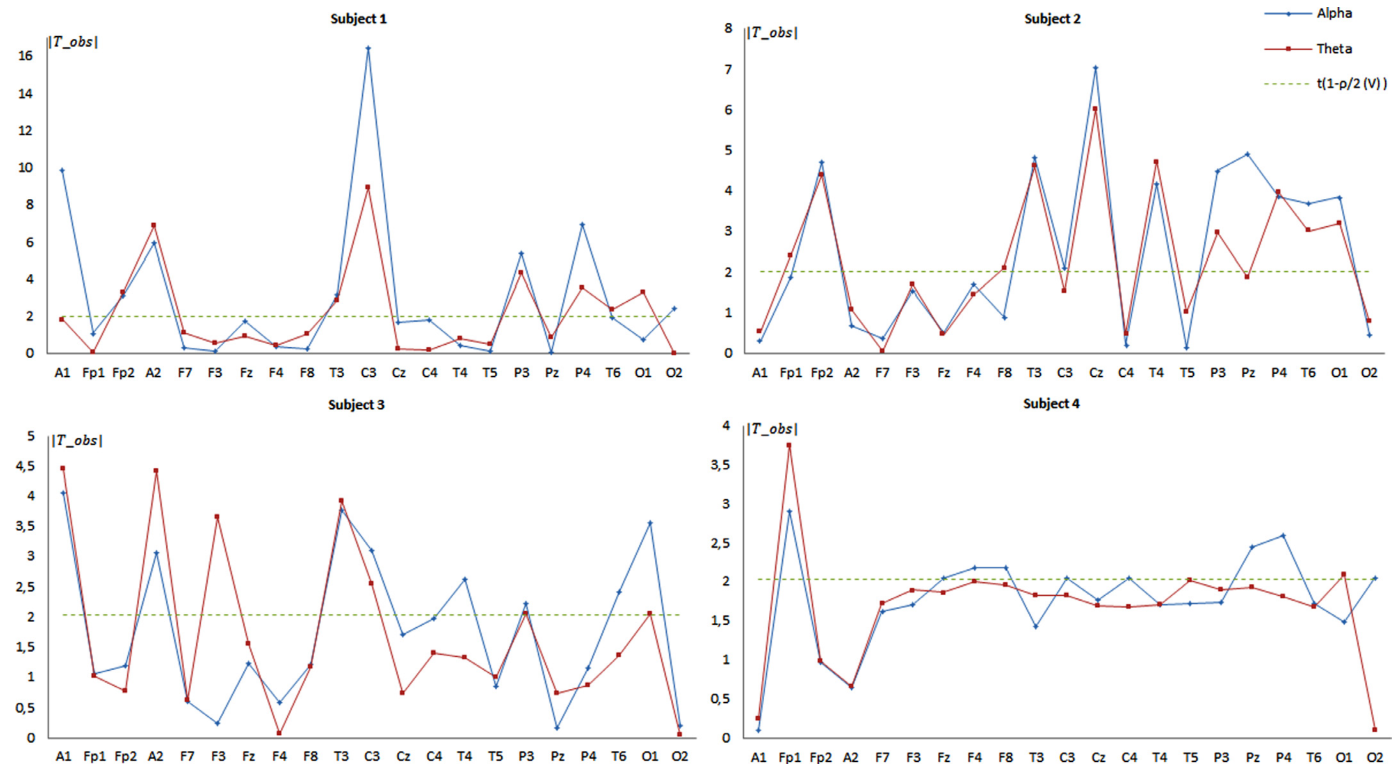

Fig. 3. Results of the workload classifier obtained on the four subjects from the first experimental protocol.

Table 2
Performance characteristic parameters of the developed classifier (second experimental protocol).

Subjects	θ			α		
	TPR (%)	SPC (%)	ACC (%)	TPR (%)	SPC (%)	ACC (%)
S1	50	75	62.5	66.6	100	75
S2	100	100	100	80	75	87.5
S3	50	100	62.5	100	75	87.5
S4	75	100	87.5	60	100	75
S5	100	85.7	87.5	42	100	50
S6	50	100	87.5	100	100	100
S7	50	100	87.5	100	100	100
S8	50	83.3	75	100	100	100
Average	65.62	70.5	81.25	69.8	93.75	84.37

calculated according to the two equations displayed in (8) and (9).

$$TPR = \frac{TP}{TP + FN} \tag{8}$$

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

One may note that basing on the absolute value of the statistic $|T_{obs}|$ we only obtain information of the decision made on changes in the workload level. In order to discover the nature of change (increase or decrease) in the level of workload, the use of raw values of the designed statistic is required. The sign of the statistic value is directly related to the increase or decrease in the workload level which is, in turn, directly related to the difficulty level of the cognitive task. In order to highlight and illustrate the behavior of these changes, we present in Figs. 5 and 6 the mean powers in θ and α bands complemented by the topographical distributions.

Results presented in these figures represent the average, across all subjects from the second experimental protocol, of the mean power of θ and α while answering questions of two different difficulty levels (l_1 and l_3). In Fig. 5, the x-axis reports the different frequencies composing the θ band. The y-axis indicates the power spectrum. Curves in each plot correspond to the power spectra of each electrode used in this experimental protocol. Results show a significant decrease in powers when increasing the difficulty of the cognitive task. This decrease is well illustrated on electrode Oz. In Fig. 6, the x-axis reports the different frequencies of the α band. For the y-axis, it presents the power spectrum. The reader can clearly note the decrease of the spectrum powers on all electrodes when moving from the cognitive task with low difficulty level to another with high difficulty level.

6. Discussion

The current study proposes a statistical methodology to classify brain workload levels during a cognitive task. The presented approach relies on (1) computing the power spectral density of the EEG signal using the Welch periodogram (2) comparing spectrums using a statistical test based on the Student's law. Through performance characteristic parameters; we showed that it provides a reasonable compromise between sensitivity and specificity. Through the experimental protocols, the averaged accuracy of classification in the case of the θ band reached 79.31% while in the case of α band it was 78.49%. Regarding these rates of correct classifications, one can note that they are significantly superior than those reported by Röy et al. (2016) and Brouwer et al. (2012). Through an experimental protocol based on the n-back test and using an SVM-based approach, Brouwer demonstrated 80% of classification while using the α band, and 72% while using the θ band. Based on the same cognitive task as Brouwer, Röy studied features of mental fatigue. This study demonstrated that the α band in the parieto-occipital lobe well describes changes in individuals' workload and mental fatigue.

In line with other literature studies, results presented by Otmani, Pebayle, Roge, and Muzet (2005) and Papadeli et al. (2006), highlight that the vigilance decrease is

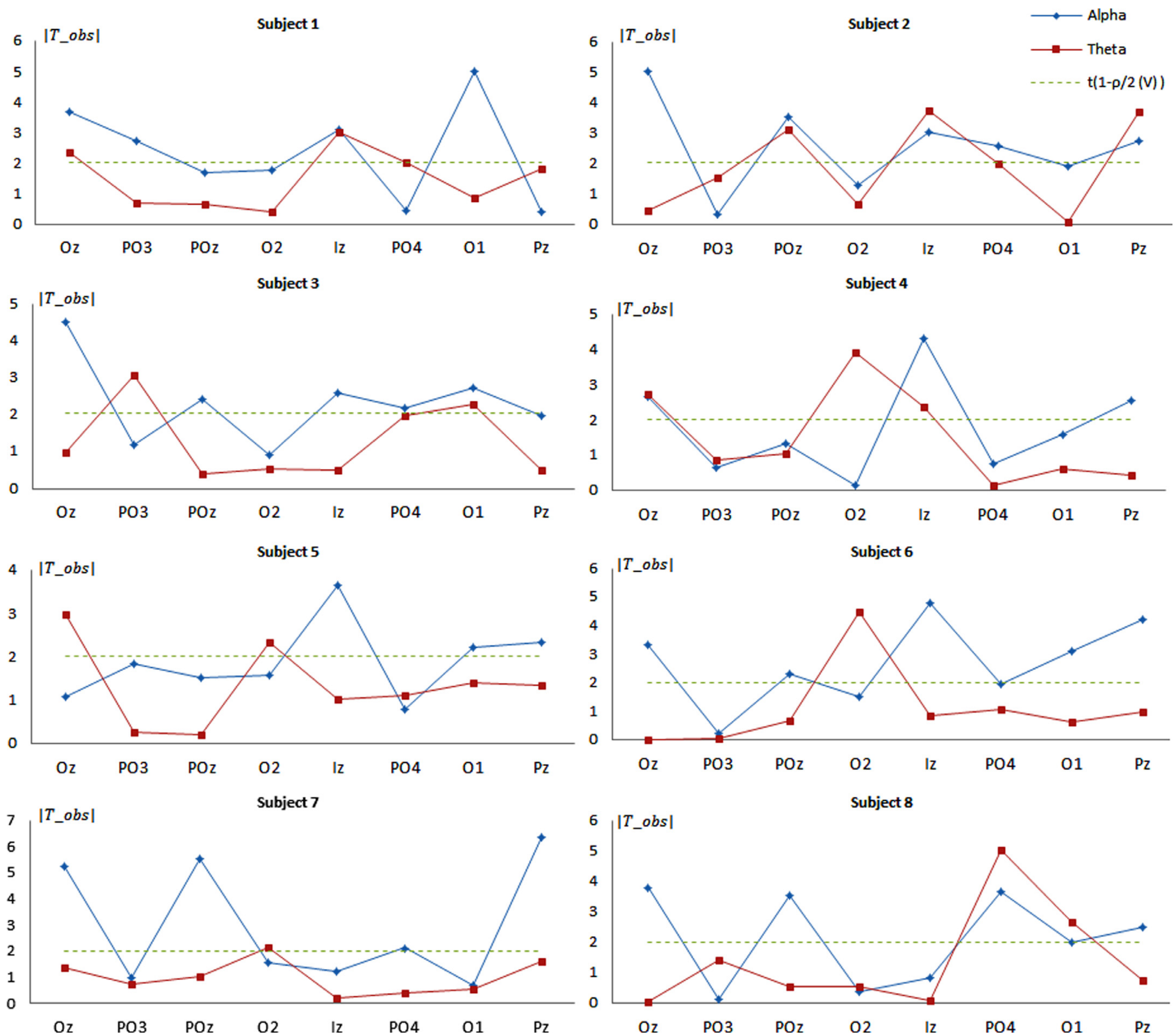

Fig. 4. Results of the workload classifier obtained on the eight subjects from the second experimental protocol.

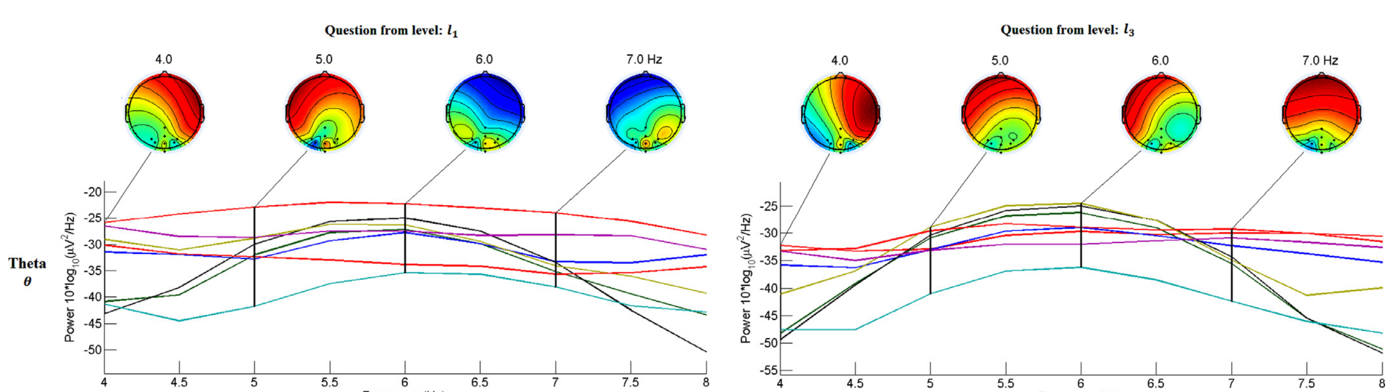

Fig. 5. Mean powers comparison of the θ activity (across participants from the second experimental protocol). Left: case of task with low workload level (l_1). Right: case of task with high workload level (l_3).

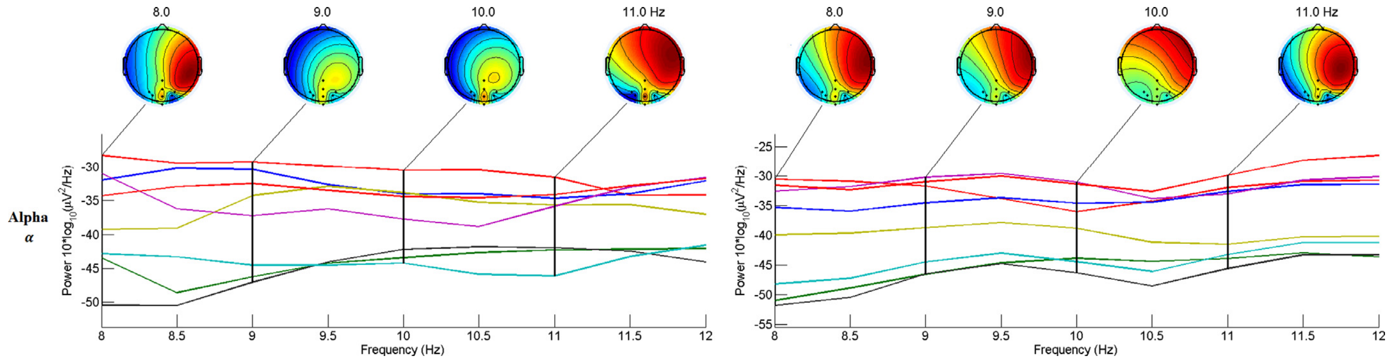

Fig. 6. Mean powers comparison of the α activity (across participants from the second experimental protocol). Left: case of task with low workload level (l_1). Right: case of task with high workload level (l_3).

characterized by an increase in θ band activity. Comparing our results to those of Otmani and Papadelis, we can deduce that the θ activity decrease, obtained in our results when increasing the difficulty level of the cognitive task, is an indicator of the subject's mental engagement and effort increase. In the case of the α rhythm, obtained results show that these oscillations are very sensitive to the increase of the task difficulty level. Comparing these results to those of Kaida, Åkerstedt, Kecklund, Nilsson, and Axelsson (2007) and Huang, Jung, and Makeig (2009), which revealed that an increase in the α band is particularly a characteristic of a reduced vigilance, we deduce that the decrease in α activity when increasing the cognitive difficulty is another indicator of the subject mental effort and vigilance increase. Comparing results of the classifier using θ and α rhythms shows high sensitivity in α wave to changes of the cognitive task difficulty levels. We propose that this sensitivity in α band, located especially in the posterior brain areas in both the two experimental protocols, reflects engagements and disengagements of the visual stream. This is due to the fact that the two cognitive tasks that we used in our experimental protocols induce high visual attention demands. This hypothesis is consistent with the results of Worden, Foxe, Wang, and Simpson (2000) which demonstrate that visual attention reduces (or even suppresses) the α activity.

Besides studying changes of brain workload when changing levels of the cognitive task, we were able to localize the brain regions related to these changes in workload. Our obtained results in this context correlate with Sakai et al. (1998) findings which identify related areas to the learning process. Based on fMRI data, Sakai identified four regions that were constantly active. These areas properly match to the brain areas identified by our classifier as areas of cognitive load changes when using data from the first protocol. Areas with prominent activation include the medial premotor cortex, the lateral prefrontal cortex, the medial parietal cortex and the lateral parietal cortex. Activation in the lateral parietal brain area corresponds to Wernicke area which is essential to language interpretation, reading ability, performance of the mathematical operations and solving logical problems. By comparing our results to those obtained with more efficient neuroimaging methods (fMRI), we demonstrate that basing only on EEG data the designed brain workload classifier performs well in identifying brain areas which interact during a cognitive task performing.

At last, some limitations to this study need to be discussed. In the first experimental protocol we induced cognitive tasks with two different difficulty levels. In fact, in the first test from this protocol, matrices sizes were $(m, n) \in \{1, 2, 3\}^2$ and composed of elements of 1 digit. This implies that the user should perform products of numbers of 1 digit and additions of numbers consisted of 2 digits at last. In the second test we introduced matrices of sizes $(m, n) \in \{4, 5, 6\}^2$ and composed of elements of 2 digits. This requires increasing the number of operations which need performing products of numbers of 2 digits and additions of numbers consisted of 3 digits at least. This experimental paradigm has generated among subjects mental efforts with averages completely distinguished. Hence the classifier efficiency. In the second experimental protocol, questions to answer were chosen so as to induce three difficulty levels (l_1, l_2 and l_3). These levels were substantially similar in pairs. Indeed, in the l_1 level, subjects were asked to complete the missing element in a numerical series. In the l_2 level, we asked subjects to solve problems like conversion of time units. Finally, in the l_3 level, questions that require problem solving using different arithmetic operations were introduced. This experimental paradigm yielded better classification rates only in the case of levels l_1 and l_3. This brings into focus the limitation of the classifier to distinguish levels of workload resulting from cognitive tasks with difficulty levels largely different. Besides this limitation, using the introduced classifier presents a robust and optimal tool to assess and detect mental fatigue in learning processes. This could be achieved by applying the classifier using a reference epoch computed when a learner is in an alert state.

7. Conclusion

In this paper we present a classifier of the brain workload changes to assess learning processes. This classifier is based on the power spectral density analysis using the Welch periodogram and the Student's law. EEG powers are calculated in the θ and α bands especially in the posterior brain areas. Results obtained from the two used experimental protocols show that increasing the difficulty of the cognitive task reduces the power spectral density of θ and α bands in the specified areas. The performance assessment, using the performance characteristic parameters evaluation method, of the designed classifier reveals that the classifier performs well in brain workload distinguishing levels. This work shows that θ and α powers could distinguish cognitive load levels in two different cognitive tasks.

In relation to expert and intelligent systems design, the theoretical contributions made by this work act particularly at the level of decisional autonomy of these systems. In the context of distance learning and intelligent tutoring systems the distinguishing the user's mental effort levels allowed us, through a work in progress, to design a new architecture of these intelligent systems. Unlike the common ITS architecture based on the four classic modules, we have exploited these theoretical contributions to add a fifth module to this architecture namely the brain workload estimation module. In this module we incorporated each of the brain signals measurement and the mental effort estimation based on the presented approach. The aim of such a module is to make intelligent decisions about the learner's mental fatigue. Based on a multi-agent system, the estimation and decision result on cognitive load levels is exploited by a "Planning" agent at the pedagogical module. This planning consists in selecting the content likely to be adequate for the current learner's mental state. According to the rules given at the pedagogical module and according to the content structuring at the domain module, the adaptation of the content can be illustrated for example by the proposal of indicators of answers in the case where the learner engages in answering an exercise with a significant mental fatigue. Moreover, we expect extending these adaptations for applications in the context of road safety. In such a context, studying the driver's brain activity represents a new promising aspect of intelligence for designing new expert systems for cars computers. At this point, the presented classifier can be adapted to make estimations of the driver's mental fatigue as well as falling asleep. For such an adaptation, this classifier can be applied by considering a reference mental state in which the driver is considered in rest state. The estimation of the mental fatigue instants must be performed based on time windows during which this classifier is applied to make the comparison to the reference state in order to make a decision. In these two contexts, i.e., learning and driving, we plan to extend the present work and exploit the obtained results to study effects of some factors and aspects such as stress, stage fright, immersion, etc.

Acknowledgments

The authors are grateful to students of both the Mohammed First University, Morocco and the Versailles University, France who accepted to participate in our experimentations. Our thanks are also addressed to the both the Laboratoire d'Ingénierie des Systèmes de Versailles, France, and the Centre Hospitalier Universitaire Mohammed VI, Oujda, Maroc for their supports.

References

Anderson, J. R., Betts, S., Ferris, J. L., & Fincham, J. M. (2012). Tracking children's mental states while solving algebra equations. Human Brain Mapping, *33*, 2650–2665.
Andreassi, J. L. (1995). *Psychophysiology: Human behavior and physiological response*. US: Lawrence Erlbaum Associates, Inc..
Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (2001). Gamma, alpha, delta, and theta oscillations govern cognitive processes. International Journal of Psychophysiology, *39*, 241–248.
Bell, C. J., Shenoy, P., Chalodhorn, R., & Rao, R. P. N. (2008). Control of a humanoid robot by a noninvasive brain–computer interface in humans. Journal of Neural Engineering, *5*, 214–220.
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series. B, Methodology, *57*, 289–300.
Brouwer, A.-M., Hogervorst, van Erp, M. A., Heffelaar, J. B. F., Zimmerman, T., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. Journal of Neural Engineering, 9.
Cutrell, E., & Tan, D. (2008). BCI for passive input in HCI. In *Proc. CHI conf. on human factors in computing systems* (pp. 1–3).
Dijksterhuis, C., de Waard, D., Brookhuis, K. A., Mulder, B. L. J. M., & de Jong, R. (2013). Classifying visuomotor workload in a driving simulator using subject specific spatial brain patterns. Frontiers in Neuroscience, 7.
Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and Clinical Neurophysiology, *70*, 510–523.
Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomics Science, *4*, 113–131.
Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology, *143*, 29–36.
Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139–183.
Healey, J. A., & Picard, R. W. (2005). Detecting Stress during real-world driving tasks using physiological sensors. IEEE Transactions on Intelligent Transportation Systems, *6*, 156–166.
Holm, A., Lukander, K., Korpela, J., Sallinen, M., & Müller, K. M. I. (2009). Estimating brain load from the EEG. The Scientific World Journal, *9*, 639–651.
Huang, K.-C., Huang, T.-Y., Chuang, C.-H., King, J.-T., Wang, Y.-K., Lin, C.-T., et al. (2016). An EEG-based fatigue detection and mitigation system. International Journal of Neural Systems, 26.
Huang, R.-S., Jung, T.-P., & Makeig, S. (2009). Tonic changes in EEG power spectra during simulated driving. In *Proc. conf. on facilities and information technology* (pp. 394–403).
Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. Psychophysiology , *37*, 163–178.
Kaida, K., Åkerstedt, T., Kecklund, G., Nilsson, J. P., & Axelsson, J. (2007). Use of subjective and physiological indicators of sleepiness to predict performance during a vigilance task. Industrial Health, *45*, 520–526.
Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. Brain Research Reviews, *29*, 169–195.
Leeb, R., Perdikis, S., Tonin, L., Biasiucci, A., Tavella, M., Creatura, M., et al. (2013). Transferring brain–computer interfaces beyond the laboratory: Successful application control for motor-disabled users. Artificial Intelligence in Medicine, *59*, 121–132.
Mühl, C., Jeunet, C., & Lotte, F. (2014). EEG-based workload estimation across affective contexts. Frontiers in Neuroscience, 8.
Otmani, S., Pebayle, T., Roge, J., & Muzet, A. (2005). Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. Physiology and Behavior, *84*, 715–724.
Pal, N. R., Chuang, C. Y., Ko, L. W., et al. (2008). EURASIP J. Adv. Signal Process. 2008: 519480. https://link.springer.com/article/10.1155/2008/519480#citeasSignal.
Papadelis, C., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Koufogiannis, D., Bekiaris, E., & Maglaveras, N. (2006). Indicators of Sleepiness in an ambulatory EEG study of night driving. In *Proc. conf. on engineering in medicine and biology society* (pp. 6201–6204).
Pires, G., Nunes, U., & Castelo-Branco, M. (2012). Evaluation of brain-computer interfaces in accessing computer and other devices by people with severe motor impairments. Procedia Computer Science, *14*, 283–292.
Postma, M. A., Schellekens, J. M. H., Hanson, E. K. S., & Hoogeboom, P. (2005). Fz theta divided by Pz alpha as an index of task load during a PC-based air traffic control simulation. In D. de Waard, K. A. Brookhuis, R. van Egmonds, & T. Boersema (Eds.), *Human factors in design, safety and management* (pp. 465–469). Maastricht: Shaker Publishing.
Putze, F., Jarvis, J.-P., & Schultz, T. (2010). Multimodal recognition of cognitive workload for multitasking in the car. In *Proc. 20th conf. on pattern recognition* (pp. 3748–3751).
Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., et al. (2010). OpenViBE: An open-source software platform to design, test, and use brain—computer interfaces in real and virtual environments. Presence: Teleoperators and Virtual Environments, *19*, 35–53.
Romo Vázquez, R., Vélez-Pérez, H., Ranta, R., Louis Dorr, V., Maquin, D., & Maillard, L. (2012). Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling. Biomedical Signal Processing and Control, *7*, 389–400.
Röy, R. N., Charbonnier, S., Campagne, A., & Bonnet, S. (2016). Efficient mental workload estimation using task-independent EEG features. Journal of Neural Engineering, 13.
Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., & Pütz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor sequence learning. Journal of Neuroscience, *18*, 1827–1840.
Spüler, M., Walter, C., Rosenstiel, W., Gerjets, P., Moeller, K., & Klein, E. (2016). EEG-based prediction of cognitive workload induced by arithmetic: A step towards online adaptation in numerical learning. *ZDM Mathematics Education, 47*, 267–278.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, *12*, 257–285.
van Dinteren, R., Arns, M., Jongsma, M. L. A., & Kessels, R. P. C. (2014). Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. Frontiers in Aging Neuroscience, 6.
van Gog, T., Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. Computers in Human Behavior, *25*, 325–331.
Wang, X.-Y., Jin, J., Zhang, Y., & Wang, B. (2014). Brain control: human-computer integration control based on brain-computer interface. Acta Automatica Sinica, *39*, 208–221.
Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Transactions on Audio and Electroacoustics, *15*, 70–73.
Wobrock, D., Frey, J., Graeff, D., de la Rivière, J.-B., Castet, J., & Lotte, F. (2015). Continuous mental effort evaluation during 3D object manipulation tasks based on brain and physiological signals. In *Proc. 15th int. conf. human-computer interaction* (pp. 472–487).
Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. Clinical Neurophysiology, *113*, 767–791.
Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electro-encephalography increases over occipital cortex. Journal of Neuroscience, 20.
Yasui, Y. (2009). A brainwave signal measurement and data processing technique for daily life applications. Journal of Physiological Anthropology, *28*, 145–150.

Zammouri, A., Aitmoussa, A., Chevallier, S., & Monacelli, E. (2015). Intelligent ocular artifacts removal in a non-invasive single channel EEG recording. In *Proc. conf. on intelligent systems and computer vision* (pp. 1–5).

Zammouri, A., & Ait Moussa, A. (2016). Effects of learning task difficulties on the prefrontal brain area. *Research Journal of Applied Sciences, Engineering and Technology, 13*, 232–236.

Zammouri, A., Chraa-Mesbahi, S., Ait Moussa, A., Zerouali, S., Sahnoun, M., Tairi, A., et al. (2017). Brain waves-based index for workload estimation and mental effort engagement recognition. *Journal of Physics: Conference Series, 904*(1), 012008.

Zhou, W., & Gotman, J. (2009). Automatic removal of eye movement artifacts from the EEG using ICA and the dipole model. Progress in Natural Science, *19*, 1165–1170.