

Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques

Hafeez Ullah Amin · Aamir Saeed Malik · Rana Fayyaz Ahmad ·
Nasreen Badruddin · Nidal Kamel · Muhammad Hussain ·
Weng-Tink Chooi

Received: 10 June 2014 / Accepted: 29 January 2015 / Published online: 4 February 2015
© Australasian College of Physical Scientists and Engineers in Medicine 2015

Abstract This paper describes a discrete wavelet transform-based feature extraction scheme for the classification of EEG signals. In this scheme, the discrete wavelet transform is applied on EEG signals and the relative wavelet energy is calculated in terms of detailed coefficients and the approximation coefficients of the last decomposition level. The extracted relative wavelet energy features are passed to classifiers for the classification purpose. The EEG dataset employed for the validation of the proposed method consisted of two classes: (1) the EEG signals recorded during the complex cognitive task—Raven’s advance progressive metric test and (2) the EEG signals recorded in rest condition—eyes open. The performance of four different classifiers was evaluated with four performance measures, i.e., accuracy, sensitivity, specificity and precision values. The accuracy was achieved above 98 % by the support vector machine, multi-layer perceptron and the K-nearest neighbor classifiers with approximation (A4) and detailed

coefficients (D4), which represent the frequency range of 0.53–3.06 and 3.06–6.12 Hz, respectively. The findings of this study demonstrated that the proposed feature extraction approach has the potential to classify the EEG signals recorded during a complex cognitive task by achieving a high accuracy rate.

Keywords Discrete wavelet transform (DWT) · Machine learning classifiers · Electroencephalography (EEG) · Cognitive task

Introduction

Electroencephalography (EEG) is a neuroimaging technique for recording the brain’s electrical potentials, which are commonly used to study the dynamics of neural information processing in the brain, and diagnose brain disorders and cognitive processes. Large amounts of EEG data are recorded and it is not possible to analyze EEG data visually [1]. Therefore, there is a strong demand to extract relevant information from EEG recordings for the proper evaluation and understanding of the desired cognitive processes. The main steps in the process of extracting relevant information from EEG recordings include pre-processing, feature extraction and classification [2]. Extracting relevant features is among the most critical and significant steps for EEG data classification. The reason behind this is that the feature extraction step has a direct impact on the systems classification performance [3]. If the extracted features are not expressive for a certain problem, then the classification performance will not be satisfactory. In such a case, the classification method may be highly optimal for the problem but due to inadequate features, the method may not provide good classification results. Hence,

H. U. Amin · A. S. Malik (✉) · R. F. Ahmad · N. Badruddin ·
N. Kamel
Department of Electrical & Electronic Engineering, Centre for
Intelligent Signal & Imaging Research (CISIR), Universiti
Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh,
Perak, Malaysia
e-mail: aamir_saeed@petronas.com.my

H. U. Amin
e-mail: hafeezullahamin@gmail.com

M. Hussain
Department of Computer Science, College of Computer and
Information Sciences, King Saud University, Riyadh 12372,
Saudi Arabia

W.-T. Chooi
Advanced Medical and Dental Institute (AMDI), Universiti
Sains Malaysia, Kepala Batas 13200, Penang, Malaysia

extracting suitable features from EEG signals to get high classification performance is mandatory.

Recently, a multi-disciplinary research area—brain computer interface (BCI), involving researchers from neuropsychology, engineering, computer science, mathematics and neuroscience attracted a lot of interest as it has the potential to provide control capabilities and communication to people facing severe motor disabilities. BCI is a system that enables a physically disabled subject to utilize brain signals to control a device without using any muscle activity. In other words, it uses the brain signals and communicates with the external devices for control. Most of the research in BCI and related research has been performed using electroencephalogram (EEG) signals. Any good practical implementation of the BCI system demands an efficient brain signal processing scheme that could extract features and perform classification [4]. Several methods have been reported for feature extraction, which include time domain, frequency domain, and wavelet transform- (WT) based features [3]. However, WT-based analysis is highly effective, because it deals better with the non-stationary behavior of EEG signals than other methods. Wavelet-based features, including wavelet entropy [5], wavelet coefficients [2], and wavelet statistical features (mean, median, and standard deviations) have been reported for normal EEG analysis as well as in clinical applications [6, 7]. Details on the performance of time domain, frequency domain and wavelet-based techniques employed in EEG classification for cognitive tasks and/or BCI applications are provided in the related work section and the classification accuracy of these techniques are provided in the discussion section. Most of these studies have reported good results in discriminating cognitive tasks of different workloads in simulated and/or real EEG recordings [8]. Hence, the experimental design used in these studies could not find dynamics in the performance of a participant in a unique task with a constant cognitive workload. Therefore, this study presents EEG signal classification in a cognitive task with a constant workload from a baseline task using wavelet-based feature extraction and machine learning classifiers.

The purpose of this study has been to extract suitable wavelet-based features (relative wavelet energy) for the classification of EEG signals. The EEG dataset used for the validation of the proposed method consisted of two classes—EEG recordings in a complex cognitive task (class 1), and EEG recordings in a rest condition—eyes open (class 2). The paper is structured as follows: “[Related work](#)” section reviews the related previous studies of feature extraction methods, “[Materials and methods](#)” section describes the materials and methods, “[Experimental results and discussion](#)” section presents the results and discussion, and “[Conclusion](#)” section concludes the paper.

Related work

In literature, the time domain, frequency domain, and wavelet-based feature extraction techniques for classification of EEG signals have been reported [9–11]. These techniques use the time and frequency domain features in the classification models to determine the optimal feature set and combine with classifiers that gives the highest classification performance. Here, we present the recent related work of the time domain, frequency domain and wavelet-based feature extraction methods for classification of EEG signals in cognitive tasks. Time domain features mainly include sample entropy [12], approximate entropy [12], permutation entropy, fractal dimension, Hjorth parameters [13], Hurst component [10], and Lyapunov exponent [14]. Frequency domain features include EEG absolute power, relative power and power ratio in different frequency bands [15]. The time–frequency analysis include wavelet-based feature extraction and stockwell transform [16]. Hariharan et al. [16] have used the stockwell transform for feature extraction and the support vector machine (SVM) for classification of EEG signals of different cognitive tasks. The authors have reported classification accuracy between 84.72 and 98.95 %. Noshadi et al. [17] have used empirical mode decomposition and both time and frequency domain features for cognitive task classification. The authors have employed linear classifiers (k-nearest neighbor and linear discriminant analysis) and have reported 97.78 % classification accuracy. Guo et al. [8] have used weighted SVM with immune feature and classified cognitive tasks with 85.4–97.5 % accuracy. Zhang et al. [18] have reported 72.4–76.4 % classification accuracy using high frequency power and Fischer’s discriminant classifier for EEG classification in cognitive tasks. Hosni et al. [19] utilized the EEG power feature and SVM classifier with a radial basis function (RBF) kernel, and have classified three cognitive tasks with 70 % accuracy. Xue et al. [20] have used the wavelet packet transform for feature extraction with the RBF classifier, and have achieved 85.3 % accuracy. Zhiwei and Minfen [21] have used the wavelet pack entropy feature and SVM classifier and have shown 87.5–93 % accuracy for discriminating a baseline task from a cognitive task. The above cited studies have used a database of seven subjects who performed five tasks—baseline (eyes open) task, multiplication, visual counting, mental letter composing and geometric object rotation. The database was originally reported by Keirn and Aunon [22] at Colorado State University. The database consisted of only seven subjects and the experimental cognitive tasks were simple. Further, the majority of the studies utilized very few subjects for classification; for example, Zhiwei and Minfen [21] have used only two subjects’ data while Nai-Jen and Palaniappan [23] have

used only four subjects' data. Additionally, the database had a variable cognitive load in different tasks.

Many other studies have worked on EEG classification in cognitive tasks using different databases recorded by the authors or adopted from past studies. Such as, Lin and Hsieh [24] have classified cognitive tasks using EEG power features with neural network classifier and have reported 78.31 % accuracy. Rodriguez-Bermudez et al. [25] have employed time, frequency and wavelet-based features with the SVM classifier and reported 67.96–80.71 % accuracy. This study has used four subjects' EEG data and a linear classifier for discrimination of cognitive tasks. Karkare et al. [26] have used a scaling exponent and classified two groups who performed complex cognitive tasks using an artificial neural network, and have reported the classification performance at over 80 % accuracy. These studies reported low classification accuracy and most of them have used non-linear classifiers, such as neural network and kernel-based SVM, which are time consuming in building models for classification. The standard psychometric cognitive task, e.g., Raven the progressive metric test, has been reported by Jahidin et al. [27]. They have utilized the EEG power feature with the neural network classifier and achieved the 88.89 % accuracy. However, this study classified within the group classification for high and low cognitive processes. From the literature, we have found a gap for this study, i.e., efficient feature extraction and classification for EEG signals for an offline dataset as well as applicable for online EEG applications.

Materials and methods

This section describes the details of the materials and methods used during this study, which include the experimental tasks, dataset description, discrete wavelet transform and relative energy computation, description of classifiers and the discussion of their performance parameters.

Participants

All of the eight healthy participants were graduate students in Universiti Teknologi PETRONAS. They participated voluntarily in this study. All of them were male, right handed and aged between 24 and 32 years (28.6 ± 4.20) [28]. At the time of the experiment, they were free from any medication, drugs, neurological disorder, or head injury that may have affected the experimental results. They had normal or corrected to normal vision. Previously, they had not experienced the cognitive task used in this study.

Consent form and ethics approval

This research study was approved by the Research Coordination Committee of Universiti Teknologi PETRONAS, Malaysia [28]. All the participants had signed the informed consent form before starting the experiment. The consent document had a brief description of this research study concerning humans.

Experimental tasks

Eyes open task

In this task, there was no cognitive task to be performed. The participants were instructed to sit relaxed and try to think of nothing in particular. To maintain the concentration of the participants, they were asked to focus their attention on a point displayed at the center of a computer screen during the EEG recording. The EEG recording of this task was used as a baseline signal.

Raven's advance progressive metric (RAPM)

RAPM is a standard psychometric test used to measure the intellectual ability. It consists of two sets (I&II). Set-I is used for practice, which contains 12 problems; Set-II consists of 36 problems used to measure the general cognitive ability. Each problem is a diagrammatic structure with some missing information and with eight multiple choices to complete the diagram's missing part. Each correct answer has a score of '1' and a score of '0' is assigned for each incorrect answer. The score range is 36 and the administration time for Set-I and Set-II is 10 and 40 min, respectively (for more details about the RAPM, see [29, 30]).

Dataset

The dataset consisted of eight healthy volunteers' EEG data, which were recorded while performing the RAPM test and eyes open condition [28, 31]. The details about the procedure of this RAPM, EEG data recording and pre-processing can be found in our previous studies [28, 31]. For feature extraction and classification, the dataset was organized into two classes. The class 1 represented all of the eight participants' EEG data recorded during the RAPM task and the class 2 represented all of the eight participants' EEG data recorded in the eyes open condition. In the RAPM task, each participant solved 36 problems. Thus, each participant was observed 36 times (maximum) in the RAPM task, resulting in 36 instances corresponding to a single participant. There were some un-attempted (not answered) problems with a few participants. The missing

problems in all of the participants were excluded and we were left with a total of 280 instances for class 1 (8 participants × 36 problems = 288, excluding missing problems ⇒ 280 instances). Similarly, each participant’s eyes open EEG recording (class 2) was segmented according to the corresponding numbers of the problem solved in the RAPM task (class 1). Hence, we have balanced both the classes in terms of instances for classification, i.e., 280 instances for each class.

Discrete wavelet transform (DWT)

The DWT is widely used for the time–frequency analysis of biomedical signals [2, 32], especially in an EEG signal analysis due to its non-stationary characteristics. The DWT employs extensive time windows for low frequencies and short time windows for higher frequencies, resulting in good time–frequency analysis. The DWT decomposition of a signal uses successive high pass and low pass filtering of the time series and two down samplers by 2. The high pass filter $g(n)$ is the discrete mother wavelet and the low pass filter $h(n)$ is its mirror version [33]. The mother wavelet of the Daubechies wavelet (db4) and the corresponding scaling function are shown in Fig. 1.

The output of the first high pass and low pass filters are referred to as the approximation and detailed coefficients, represented by A_1 and D_1 , respectively. The A_1 is further disintegrated and the procedure is repeated till the specified number of decomposition levels is reached (see Fig. 2) [32, 33].

The dilation function $\varphi_{j,k}(n)$ is dependent on the low pass filter, and the wavelet function $\psi_{j,k}(n)$ is follows the high pass filter, which is denoted as follows.

$$\varphi_{j,k}(n) = 2^{j/2}h(2^j n - k) \tag{1}$$

$$\psi_{j,k}(n) = 2^{j/2}g(2^j n - k) \tag{2}$$

where $n = 0, 1, 2, \dots, M-1$; $j = 0, 1, 2, \dots, J-1$; $k = 0, 1, 2, \dots, 2^j-1$; $J = \log_2(M)$; and M is the length of the signal [34].

The maximum level of decomposition is specified depending on the principal frequency components in the given signal [2]. The coefficients of the DWT are referred to as the dot product of the original time series and the designated basis functions. The approximation coefficients A_i and the detailed coefficients D_i in the i th level are denoted as [2]:

$$A_i = \frac{1}{\sqrt{M}} \sum_n x(n) \times \varphi_{j,k}(n) \tag{3}$$

$$D_i = \frac{1}{\sqrt{M}} \sum_n x(n) \times \psi_{j,k}(n) \tag{4}$$

where $k = 0, 1, 2, \dots, 2^j-1$ and M is the length of the EEG time series in the discrete points.

Relative and total wavelet sub-band energy

The wavelet energy at each decomposition level $i = 1, \dots, L$ is computed as follows:

$$E_{D_i} = \sum_{j=1}^N |D_{ij}|^2, \quad i = 1, 2, 3, \dots, L \tag{5}$$

$$E_{A_i} = \sum_{j=1}^N |A_{ij}|^2, \quad i = L \tag{6}$$

The ‘ L ’ is the maximum level of decomposition. Hence, from Eqs. 5 and 6, the total energy can be defined as:

$$E_{Total} = \left(\sum_{i=1}^L E_{D_i} + E_{A_L} \right) \tag{7}$$

The normalized energy values represent the relative wavelet energy.

$$E_r = \frac{E_j}{E_{Total}} \tag{8}$$

where $E_j = E_{D_{i=1, \dots, L}}$ or $E_{A_{i=L}}$

Fig. 1 Mother wavelet and scaling function (db4)

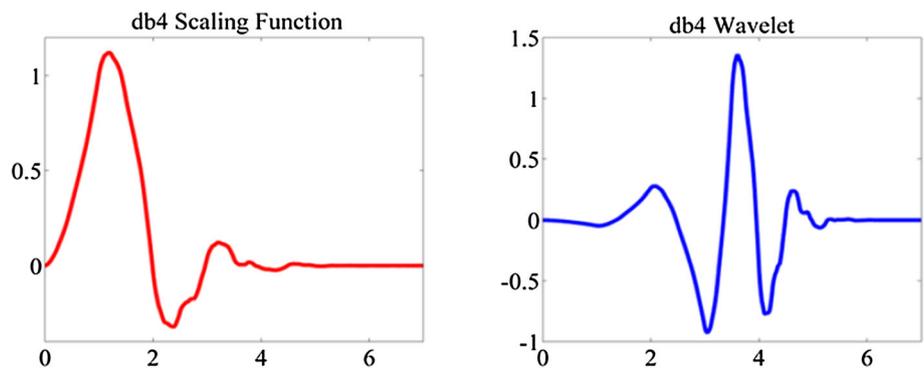
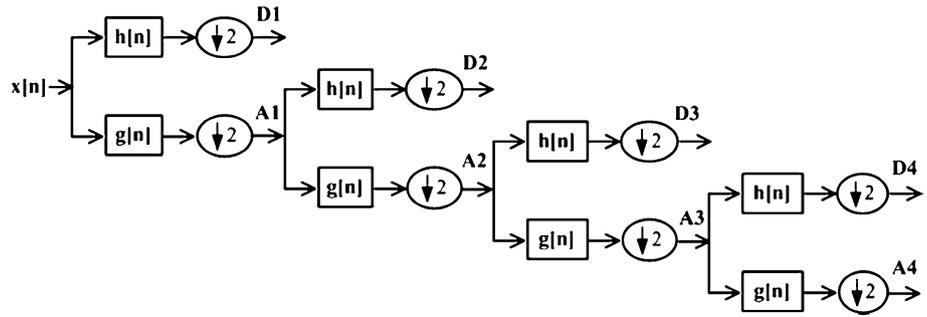


Fig. 2 DWT sub-band decomposition



Feature extraction using DWT

The relevant information extraction from raw signals is a critical step in EEG pattern classification due to its direct influence on classification performance. The illustrative representation of the proposed feature extraction scheme is presented in Fig. 3. The band pass (1–48 Hz) EEG signal was decomposed into sub-band frequencies by using the discrete wavelet transformation with the Daubechies wavelet of order 4 up to level 4. The approximate and detailed coefficients were computed (see Fig. 4 as an example). Table 1 represents one channel’s sub-band percentage relative energy and its frequency range of a single subject. The total and relative sub-band energies were computed from the extracted wavelet coefficients. The relative wavelet energy $E_{rD1}, E_{rD2}, \dots, E_{rA4}$ was calculated using Eq. 8.

The relative energy features were computed for all of the participants and all of the channels’ data. Accordingly, the feature matrix of relative energy for a single participant in each EEG task and each sub-band (detailed or approximation) became as follows:

$$\text{Relative Energy Feature Matrix}(\vec{F}_r) = [E_{rA4(280 \times 128)}] \quad (9)$$

where the number of channels was 128, number of instances in each class was 280, $D_1 \sim D_4$ and A_4 were the detailed and approximation coefficients. Accordingly, the $E_{D1(280 \times 128)}$ represented the relative energy feature matrix of the first detailed coefficients for all of the eight participants in each class.

Classification methods

A classifier is a technique that utilizes various independent variable values (features) as input and predicts the corresponding class to which the independent variable belongs [12]. In the EEG signal analysis, the features can be any kind of extracted information from the signal, such as energy, entropy, power etc. and the class can be the type of task or the stimulus used during the recording. A classifier has a number of parameters that need to be learned from training data. The learned classifier is a model of the association between the features and the classes. For example, for a given feature x of a class y , the classifier is a function f that predicts the class $y = f(x)$. After the learning, the classifier is able to predict new instances that have not been used in the training data. Thus, the performance of the classifier is tested on a different set of instances.

To demonstrate the effectiveness of the proposed feature extraction scheme in cognitive function classification, the SVM, multi-layer perceptron (MLP), K-nearest neighbor (K-NN) and Naïve Byes classifiers were used. The SVM used a kernel trick to transform the data points into a higher dimensional space and then separated them by a hyper-plane with a maximal margin. The MLP is a neural network-based method, which is commonly used for performing a different variety of detection and estimation tasks. The K-nearest neighbor works to find a testing sample’s class by the majority class of the k nearest training samples. The Naïve Bayes is a simple and efficient

Fig. 3 Proposed scheme for feature extraction and classification of EEG signal

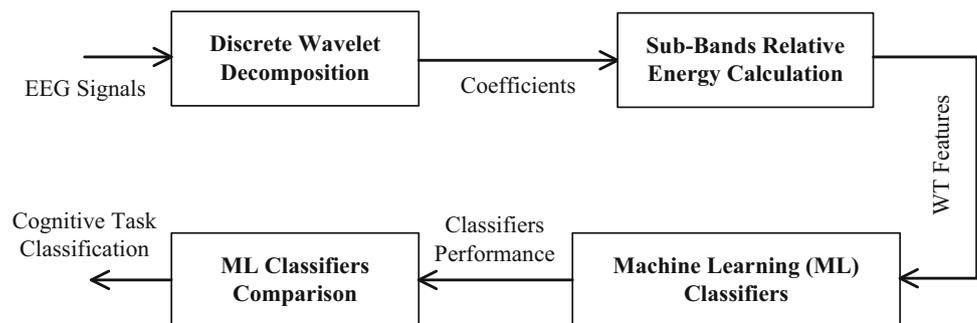


Fig. 4 Representation of the A4 and D1–D4 components of one participant's EEG signal at the F3 scalp location during a cognitive task

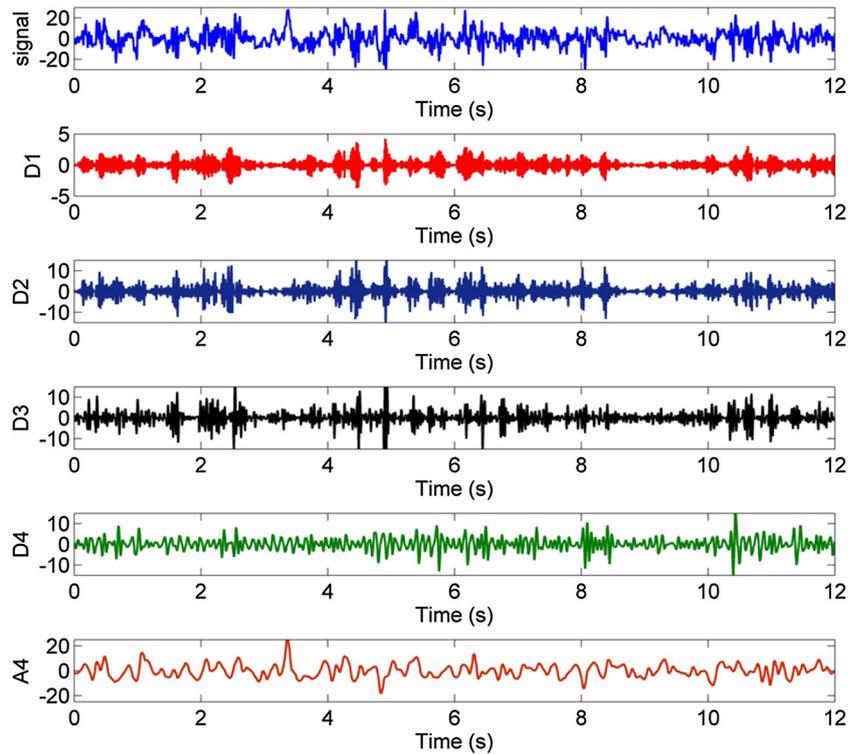


Table 1 Frontal F3 channel's sub-band percentages' relative energy and their frequency range

Levels	Wavelet energy (%)	Wavelet coefficients	Frequency bands (Hz)
1	0.48	D1	23.5–48
2	8.47	D2	12.25–23.5
3	10.70	D3	6.12–12.25
4	13.63	D4	3.06–6.12
4	70.96	A4	0.53–3.06

D1 detailed coefficient of level 1, *D2* detailed coefficient of level 2, *D3* detailed coefficient of level 3, *D4* detailed coefficient of level 4, *A4* approximation level 4

statistical method, which is based on Bayes' theorem. For more details about SVM, MLP, K-NN and Naïve Bayes (see [1, 2, 35–37]).

Experimental results and discussion

In this section, we present the experimental results for validation, and discuss them. We start with the experimental set-up used for these experiments.

Experimental setup

The classifiers were trained and tested for these extracted features using the tenfold cross validation method. The

tenfold cross validation method divided the dataset into ten subsets of equal size and used nine subsets for the classifier training and one subset for the classifier testing. This process repeated ten times, each time leaving out one of the subsets from the training, which was used for testing. This method has the advantage that it utilized all of the instances in the dataset for both training and testing. The classifiers' performances were computed using the most commonly used parameters, are the accuracy, sensitivity, specificity, precision, and Kappa statistic [38]. These parameters are defined as follows.

$$\text{Accuracy} = \frac{\text{Total no. of correctly classified instances}}{\text{Total numbers of instances}} \times 100 \quad (10)$$

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \times 100 \quad (11)$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}} \times 100 \quad (12)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \times 100 \quad (13)$$

$$\text{Kappa}(k) = \frac{(P_o - P_e^C)}{(1 - P_e^C)} \quad (14)$$

where P_o represents the probability of the overall agreement of the label assignments between the classifier and the true process, and P_e^C denotes the chance agreement over the

labels—sum of the proportion of instances assigned to a class multiplies the proportion of true labels of that class in the dataset.

Classification results

The extracted relative wavelet energy features of D1–D4 and A4 were classified using SVM with RBF kernel, MLP with five hidden layers, K-NN with $k = 1$, and Naïve Bayes classifiers for both of the EEG conditions, i.e., eyes open and cognitive task. This classification process was implemented for the extracted features from all of the decomposition levels (D1–D4, and A4). However, the classification results were not prominent in all of the decomposition levels. The highest classification performance was found in the relative energy of the approximation coefficients and detailed coefficients of level 4, which reflected the low frequency (0.53–3.06 Hz) and above low frequency (3.06–6.12 Hz) dominations in the cognitive task (see Tables 2 and 3). A representative signal of 8 s from both the experimental tasks at the F3 electrode is presented in Fig. 5. The amplitude differences can be observed in both the 0.5–3 and 3–6 Hz frequency bands of the two experimental tasks.

The SVM classifier achieved 98.75 % accuracy and the MLP and K-NN classifiers achieved 98.21 %, accuracy in the classification using the relative energy of the approximation A4 coefficients. In the detailed D4 coefficients, the SVM and MLP achieved 98.21 and 98.57 % accuracy, respectively, as shown in Tables 2 and 3. The accuracy of the Naïve Bayes classifier was also found to be above 80 %. The values of the other performance parameters,

such as sensitivity, specificity, precision and Kappa statistic were prominent. From these results, it seems that the relative wavelet energy in the low frequency band (0.53–3.06 Hz) and the above low frequency band (3.06–6.12 Hz) was a useful feature to classify the EEG brain patterns in eyes open and the complex cognitive task (i.e., RAPM).

Discussion

Comparison with existing techniques

A direct comparison of the results with the previous research in EEG signals was hard due to the variety of EEG datasets, wavelet types, decomposition levels, participants' variability, and the cognitive tasks used. However, a brief comparison with the previous related studies is presented here. The information about the dataset, feature extraction methods, cognitive tasks, machine learning algorithm and the classification performance reported in previous studies are presented in Table 4. The list of studies in Table 4 have used the time domain, frequency domain, autoregressive (AR) coefficient and/or wavelet transform-based features for EEG classification in a cognitive task as mentioned in “[Related work](#)” section. The majority of the studies have used non-linear classifiers (e.g., ANN and kernel-based SVM), which are complex in nature and time consuming to build the classification model. In the case of using very few instances in the classification as mentioned in a few studies in Table 4, it may causes the over fitting problem in classification [39]. In this work, we used 280 instances for each class in the

Table 2 Classification results of the relative wavelet energy of the level 4 approximate coefficients (A4) for the cognitive task

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC	Kappa statistic
SVM	98.75	100	97.50	97.60	0.98	0.97
MLP	98.21	100	96.40	96.60	0.98	0.96
K-NN	98.21	99.60	96.80	96.90	0.98	0.96
Naïve Bayes	83.57	75.00	92.10	90.50	0.93	0.67

SVM support vector machine, *MLP* multi-layers perceptron, *K-NN* k-nearest neighbors, *AUC* area under the receiver operating characteristic curve

Table 3 Classification results of the relative wavelet energy of the level 4 detailed coefficients (D4) for the cognitive task

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	AUC	Kappa statistic
SVM	98.21	99.60	96.80	96.90	0.98	0.96
MLP	98.57	99.60	97.60	97.60	0.98	0.97
K-NN	97.14	98.90	95.40	95.50	0.97	0.94
Naïve Bayes	83.03	84.60	81.40	82.00	0.89	0.66

SVM support vector machine, *MLP* multi-layers perceptron, *K-NN* k-nearest neighbors, *AUC* area under the receiver operating characteristic curve

classification. Such a high number of instances in classification have not been reported in cognitive task classification using EEG. Hence, we have used both the linear (e.g., K-NN) as well as the non-linear classifiers (e.g., ANN) along with the 10-cross validation scheme. The benefit of the 10-cross validation process is that all instances in the sample are used for both training and validation exactly once [39]. Therefore, in the presence of a high number of instances, the use of multiple classifiers in the present study made it comparable with previous studies in terms of classification performance. The classification results of this study in both linear and non-linear classifiers were found to be better than related studies which used similar classifiers and the same nature of the cognitive tasks.

EEG low frequencies with cognitive neuroscience perspective

The EEG low frequency bands (delta and theta) have been reported by the cognitive neuroscientists as cognitive rhythms, and have been linked with cognitive and attention demanded tasks [43–45]. Especially, the event-related potential (ERP) studies have reported the most significant findings of the delta band related to cognitive processing [44], i.e., the associations of the P300 component with the cognitive process [46]. This relationship has been widely reported in the cognitive neuroscience literature. In brief, the delta band has been considered as the primary

contributor to the P300 component of ERP [47]. Gennady et al. [48] reviewed the delta band relationship with cognitive processing and have confirmed that delta is linked with the cognitive process. Similarly, the theta rhythms are the most intensively studied in the cognitive neuroscience aimed at correlating the theta rhythms with cognitive processing [49–51]. Particularly, theta in the frontal regions is critical for attentional and cognitive processing in ERP tasks [50]. Most of these studies reported significant increase in delta and theta power in the cognitive tasks. This may be the reason, that we achieved high classification accuracy in the low frequency bands (0.53–3.06 and 3.06–6.12 Hz) for discriminating the RAPM and baseline—eyes open task. Hence, the results of this study reflect the previous studies' findings in cognitive neuroscience research.

Conclusion

This paper has presented the use of relative discrete wavelet energy along with machine learning algorithms for the classification and the quantitative analysis of spontaneous EEG signals recorded during complex cognitive task. The EEG signals were split into sub-bands using DWT with Daubechies (db4) wavelets and the sub-bands' relative energies were computed for all of the 128 channels of each subject's EEG recording. For classification, four

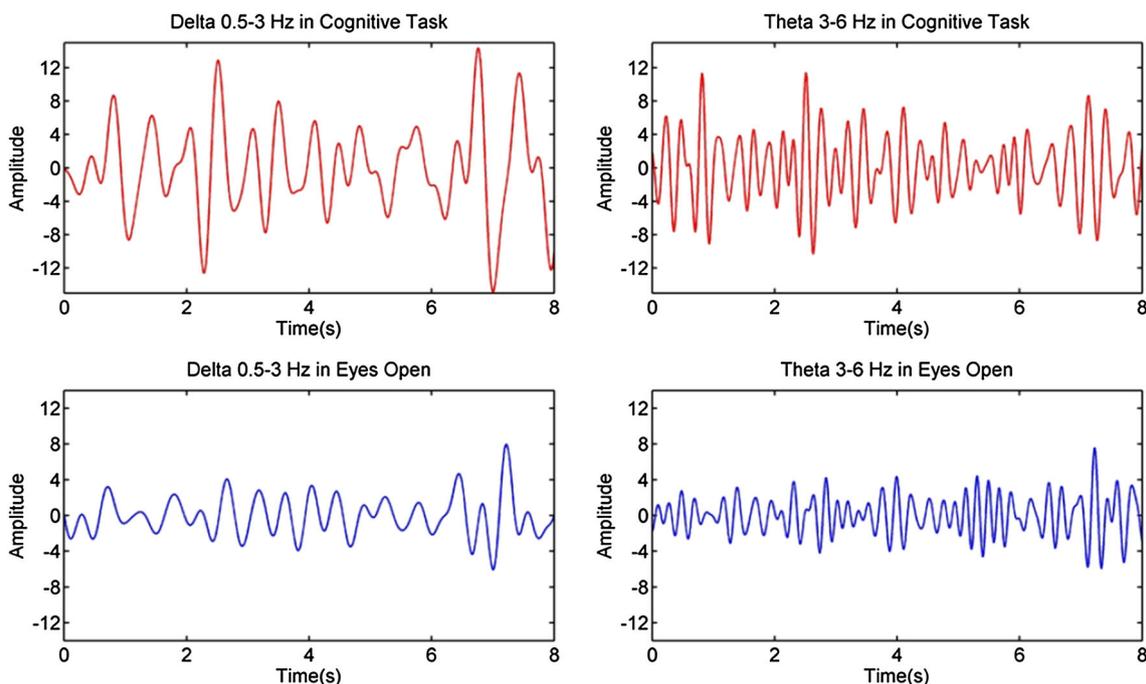


Fig. 5 Representative signal of low frequencies (delta and theta bands) at the F3 electrode position (*Red color* shows the signal of the cognitive task, and the *blue color* represents the signal of the eyes open task)

Table 4 Summary of existing feature extraction and classification techniques for EEG in cognitive tasks

Ref	Year	Subjects (instances)	EEG scalp electrodes	Cognitive task	Feature	Classifier	Accuracy (%)
[20]	2003	4	6	Five tasks as reported by Keirn and Aunon [22]	Wavelet packet	RBF network	85.30
[40]	2004	4	6	Five tasks as reported by Keirn and Aunon [22]	DWT	ANN	74.40–82.30
[4]	2006	7	6	Five tasks as reported by Keirn and Aunon [22]	AR coefficients	SVM, ANN and ELM	53.98–56.07
[21]	2007	2	6	Five tasks as reported by Keirn and Aunon [22]	Wavelet Packet Entropy up to seven levels	SVM	87.5–93.0
[19]	2007	7	6	Three tasks classified (baseline, mental letter, multiplication) from Keirn and Aunon [22]	Frequency Bands Power and AR coefficients	SVM with RBF kernel	70
[41]	2009	×	×	–	Wavelet relative energy	ANN	95.2
[42]	2009	7	6	Five tasks as reported by Keirn and Aunon [22]	Empirical mode decomposition, time and frequency domain features	ANN and LDA	87.35–91.17
[6]	2009	4	6	Five tasks (Baseline or relaxed state, multiplication, visual counting, mental letter composing geometric object rotation) as reported by Keirn and Aunon [22]	DWT with db4 up to level 5	K-NN	81.48–89.58
[24]	2009	3	8	Three cognitive tasks (imagination of left hand movement, imagination of right hand movement, and generation of words)	Power spectral density feature	ANN and SVM	65.90–68.35
[8]	2011	7	6	Five tasks as reported by Keirn and Aunon [22]	Immune feature	SVM	85.4–97.5
[27]	2014	(120)	×	Raven's progressive metric	Frequency Band power	ANN	88.89
This work	2014	8(560)	128	Baseline (eyes open), RAPM	Wavelet relative energy	SVM MLP K-NN Naïve Bayes	98.75 98.57 98.21 83.57

SVM support vector machine, MLP multi-layers perceptron, K-NN k-nearest neighbors, AUC area under the receiver operating characteristic curve, AR autoregressive, RBF radial basis function, ANN artificial neural network, LDA linear discriminant analysis, DWT Discrete wavelet transform, RAPM Raven's advance progressive metric, db4 Daubechies wavelet order 4, × unknown. The values in brackets '()' indicate number of instances and values without bracket or outside bracket denote number of subjects only in the 3rd column from side of the above table

different classifiers (SVM, MLP, K-NN and Naïve Bayes) were employed and their performance was evaluated for cognitive task discrimination. The classification results of SVM and MLP demonstrated above 98 % accuracy with features extracted using A4 (0.53–3.06 Hz) and D4 (3.06–6.12 Hz) sub-bands. The wavelet energy is a useful feature to classify the EEG signals corresponding to complex cognitive tasks, and it will be helpful for EEG classification in clinical applications, such as epilepsy, depression, and stress diagnosis as it is capable of identifying variations in non-stationary EEG signals because of

the localization characteristics of the wavelet transformation [5]. The low frequencies, especially in the range of the delta band, are perceived in cognitive neuroscience as the primary contributor to the cognitive processing. Hence, the proposed feature scheme has the clinical significance to be applied on real time EEGs in BCI applications for severe motor disabled patients to control external devices using cognitive power. This may be implemented in future work.

Acknowledgments This research work has been supported by University Research Internal Funding (URIF: 0153AA-B26), Universiti

Teknologi PETRONAS; the Fundamental Research Grant Scheme (Ref: FRGS/1/2014/TK03/UTP/02/1), Ministry of Education (MOE), Malaysia and by NSTIP strategic technologies programs, Grant number (12-INF2582-02), in the Kingdom of Saudi Arabia.

References

- Übeyli ED (2009) Combined neural network model employing wavelet coefficients for EEG signals classification. *Digit Signal Proc* 19:297–308
- Orhan U, Hekim M, Ozer M (2011) EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. *Expert Syst Appl* 38:13475–13481
- Iscan Z, Dokur Z, Demiralp T (2011) Classification of electroencephalogram signals with combined time and frequency features. *Expert Syst Appl* 38:10499–10505
- Liang N-Y, Saratchandran P, Huang G-B, Sundararajan N (2006) Classification of mental tasks from EEG signals using extreme learning machine. *Int J Neural Syst* 16:29–38
- Rosso OA, Blanco S, Yordanova J, Kolev V, Figliola A, Schürmann M et al (2001) Wavelet entropy: a new tool for analysis of short duration brain electrical signals. *J Neurosci Methods* 105:65–75
- Yazdani A, Ebrahimi T, Hoffmann U (2009) Classification of EEG signals using Dempster Shafer theory and a k-nearest neighbor classifier. In: 4th International IEEE/EMBS conference on neural engineering, pp 327–330
- Garry H, McGinley B, Jones E, Glavin M (2013) An evaluation of the effects of wavelet coefficient quantisation in transform based EEG compression. *Comput Biol Med* 43:661–669
- Guo L, Wu Y, Zhao L, Cao T, Yan W, Shen X (2011) Classification of mental task from EEG signals using immune feature weighted support vector machines. *IEEE Trans Magn* 47:866–869
- Taghizadeh-Sarabi M, Daliri MR, Niksirat KS (2014) Decoding objects of basic categories from electroencephalographic signals using wavelet transform and support vector machines. *Brain Topogr* 1–14
- Acharya UR, Sree SV, Ang PCA, Yanti R, Suri JS (2012) Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals. *Int J Neural Syst* 22:1250002
- Gandhi T, Panigrahi BK, Anand S (2011) A comparative study of wavelet families for EEG signal classification. *Neurocomputing* 74:3051–3057
- Richman JS, Moorman JR (2000) Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 278:H2039–H2049
- Vidaurre C, Krämer N, Blankertz B, Schlögl A (2009) Time domain parameters as a feature for EEG-based brain-computer interfaces. *Neural Netw* 22:1313–1319
- Übeyli ED (2010) Lyapunov exponents/probabilistic neural networks for analysis of EEG signals. *Expert Syst Appl* 37:985–992
- Thatcher RW, North D, Biver C (2005) EEG and intelligence: relations between EEG coherence, EEG phase delay and power. *Clin Neurophysiol* 116:2129–2141
- Hariharan M, Vijean V, Sindhu R, Divakar P, Saidatul A, Yaacob S (2014) Classification of mental tasks using stockwell transform. *Comput Electr Eng* 40:1741
- Noshadi S, Abootalebi V, Sadeghi MT, Shahvazian MS (2014) Selection of an efficient feature space for EEG-based mental task discrimination. *Biocybern Biomed Eng* 34:159–168
- Zhang L, He W, He C, Wang P (2010) Improving mental task classification by adding high frequency band information. *J Med Syst* 34:51–60
- Hosni SM, Gadallah ME, Bahgat SF, AbdelWahab MS (2007) Classification of EEG signals using different feature extraction techniques for mental-task BCI. In: International conference on computer engineering & systems, 2007. ICCES'07, pp 220–226
- Xue J-Z, Zhang H, Zheng C-x, Yan X-G (2003) Wavelet packet transform for feature extraction of EEG during mental tasks. In: International conference on machine learning and cybernetics, 2003, pp 360–363
- Zhiwei L, Minfen S (2007) Classification of mental task EEG signals using wavelet packet entropy and SVM. In: 8th International conference on electronic measurement and instruments, 2007. ICEMI'07, pp 3-906–3-909
- Keirn ZA, Aunon JI (1990) A new mode of communication between man and his surroundings. *IEEE Trans Biomed Eng* 37:1209–1214
- Nai-Jen H, Palaniappan R (2004) Classification of mental tasks using fixed and adaptive autoregressive models of EEG signals. In: 26th Annual international conference of the IEEE engineering in medicine and biology society, 2004. IEMBS'04, pp 507–510
- Lin C-J, Hsieh M-H (2009) Classification of mental task from EEG data using neural networks based on particle swarm optimization. *Neurocomputing* 72:1121–1130
- Rodríguez-Bermúdez G, García-Laencina PJ, Roca-González J, Roca-Dorda J (2013) Efficient feature selection and linear discrimination of EEG signals. *Neurocomputing* 115:161–165
- Karkare S, Saha G, Bhattacharya J (2009) Investigating long-range correlation properties in EEG during complex cognitive tasks. *Chaos Solitons Fractals* 42:2067–2073
- Jahidin A, Ali MM, Taib MN, Tahir NM, Yassin IM, Lias S (2014) Classification of intelligence quotient via brainwave sub-band power ratio features and artificial neural network. *Comput Methods Programs Biomed* 114:50–59
- Amin HU, Malik AS, Subhani AR, Badruddin N, Chooi W-T (2013) Dynamics of scalp potential and autonomic nerve activity during intelligence test. In: Lee M et al (eds) *Neural information processing*, vol 8226. Springer, Berlin, pp 9–16
- Raven J (2000) The Raven's progressive matrices: change and stability over culture and time. *Cogn Psychol* 41:1–48
- Kunda M, Mcgreggor K, Goel A (2012) Reasoning on the Raven's advanced progressive matrices test with iconic visual representations. In: 34th Annual conference of the cognitive science society pp 1828–1833
- Amin HU, Malik AS, Badruddin N, Chooi W-T (2013) EEG mean power and complexity analysis during complex mental task. In: ICME International conference on complex medical engineering (CME) pp 648–651
- Jahankhani P, Kodogiannis V, Revett K (2006) EEG signal classification using wavelet feature extraction and neural networks. In: IEEE John Vincent Atanasoff 2006 international symposium on modern computing, 2006. JVA'06. pp 120–124
- Subasi A (2007) EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Syst Appl* 32:1084–1093
- Gonzalez RC, Woods RE (2002) *Digital image processing*, 2nd edn. Prentice Hall, SL
- Hausfeld L, De Martino F, Bonte M, Formisano E (2012) Pattern analysis of EEG responses to speech and voice: influence of feature grouping. *Neuroimage* 59:3641–3651
- Ben-Hur A, Weston J (2010) A user's guide to support vector machines. In: *Data mining techniques for the life sciences*, Springer, New York, pp 223–239
- Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45:S199–S209
- Subasi A, Gursoy MI (2010) EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst Appl* 37:8659–8666

39. Witten IH, Frank E (2005) Data mining: practical machine learning tools and techniques. Morgan Kaufmann, Burlington
40. Daud M, Yunus J (2004) Classification of mental tasks using denoised EEG signals. In: 7th International conference on signal processing, pp 2206–2209
41. Guo L, Rivero D, Seoane JA, Pazos A (2009) Classification of EEG signals using relative wavelet energy and artificial neural networks. In: Proceedings of the first ACM/SIGEVO summit on genetic and evolutionary computation, pp 177–184
42. Diez PF, Mut V, Laciari E, Torres A, Avila E (2009) Application of the empirical mode decomposition to the extraction of features from EEG signals for mental task classification. In: Annual international conference of the IEEE engineering in medicine and biology society, 2009. EMBC 2009, pp 2579–2582
43. Dimitriadis SI, Laskaris NA, Tsirka V, Vourkas M, Micheliyannis S (2010) What does delta band tell us about cognitive processes: a mental calculation study. *Neurosci Lett* 483:11–15
44. Harper J, Malone SM, Bernat EM (2014) Theta and delta band activity explain N2 and P3 ERP component activity in a go/no-go task. *Clin Neurophysiol* 125:124–132
45. Amin H, Malik AS (2013) Human memory retention and recall processes: a review of EEG and fMRI studies. *Neurosciences* 18:330–344
46. Ergen M, Marbach S, Brand A, Başar-Eroğlu C, Demiralp T (2008) P3 and delta band responses in visual oddball paradigm in schizophrenia. *Neurosci Lett* 440:304–308
47. Demiralp T, Ademoglu A, Schürmann M, Basar-Eroglu C, Basar E (1999) Detection of P300 waves in single trials by the wavelet transform (WT). *Brain Lang* 66:108–128
48. Gennady GK (2012) EEG delta oscillations as a correlate of basic homeostatic and motivational processes. *Neurosci Biobehav Rev* 36:677–695
49. Baijal S, Srinivasan N (2010) Theta activity and meditative states: spectral changes during concentrative meditation. *Cogn Process* 11:31–38
50. Sakowitz O (2001) The selectively distributed theta system: functions. *Int J Psychophysiol* 39:197–212
51. Nigbur R, Ivanova G, Stürmer B (2011) Theta power as a marker for cognitive interference. *Clin Neurophysiol* 122:2185–2194