Subject identification from low-density EEG-recordings of resting-states: A study of feature extraction and classification

Luis Alfredo Moctezuma and Marta Molinas

Department of Engineering Cybernetics, Norwegian University of Science and Technology, 7491 Trondheim, Norway
luisalfredomoctezuma@gmail.com, marta.molinas@ntnu.no

Abstract. A new concept of low-density electroencephalograms-based (EEG) Subject identification is proposed in this paper. To that aim, EEG recordings of resting-states were analyzed with 3 different classifiers (*SVM, k-NN*, and *naive Bayes*) using Empirical Mode Decomposition (EMD) and Discrete Wavelet Transform (DWT) for feature extraction and their accuracies were estimated to compare their performances. To explore the feasibility of using fewer channels with minimum loss of accuracy, the methods were applied to a dataset of 27 Subjects (From 5 sessions of 30 instances per Subject) recorded using the EMOTIV EPOC device with 1 set of 14 channels and 4 subsets (8, 4, 2 and 1 channel) that were selected using a *greedy* algorithm. The experiments were reproduced using fewer instances each time to observe the evolution of the accuracy using both; fewer channels and fewer instances. The results of this experiments suggest that EMD compared with DWT is a more robust technique for feature extraction from brain signals to identify Subjects during resting-states, particularly when the amount of information is reduced: e.g., using *Linear SVM* and 30 instances per Subject, the accuracies obtained using 14 channels were 0.91 and 0.95, with 8 channels were 0.87 and 0.89 with EMD and DWT respectively but were reversed in favor of EMD when the number of channels was reduced to 4 channels (0.76 and 0.74), 2 (0.64 and 0.56) and 1 channel (0.46 and 0.31). The general observed trend is that, *Linear SVM* exhibits higher accuracy rates using high-density EEG (0.91 with 14 channels) while *Gaussian naive Bayes* exhibits better accuracies when using low-density EEG in comparison with the other classifiers (With EMD 0.88, 0.81, 0.76 and 0.61 respectively for 8, 4, 2 and 1 channel). The findings of these experiments reveal an important insight for continuing the exploration of low-density EEG for Subject identification.

Keywords: Biometric security, Subject identification, Electroencephalograms (EEG), Resting-states, Empirical Mode Decomposition (EMD), Discrete Wavelet Transform (DWT)

1 Introduction

To protect places and/or information where privileges are required, organizations use security systems. To achieve this, different measures have been proposed, ranging from security-guards/smart-cards to fingerprint/face-recognition [1, 2].

The use of security systems has been increasing not only in organizations but also in low-cost portable devices (e.g., mobile phones, tablets and personal computers). Due to the increasing vulnerabilities to skip the authentication and authorization process of current traditional/biometric security systems [2], there is a growing interest in exploring new biometric measures. With this trend, the use of brain signals to create biometric markers using different neuro-paradigms also has emerged as a robust alternative to the above mentioned vulnerabilities. Brain signals can be used as a basis for the design of biometric markers since they satisfy the requirements of *universality, permanence, collectability, performance, acceptability, and circumvention* [1]. Brain signals are more reliable and secure because biometric markers obtained from EEG-recordings from human brain activity will be almost impossible to duplicate since the brain is highly individual [3].

To promote the concept of “low-cost” affordable devices to record brain signals using different neuro-paradigms, a popular/non-invasive technique using Electroencephalography (EEG) is the well known Brain-Computer Interface (BCI).

Empirical Mode Decomposition (EMD) [4] and Discrete Wavelet Transform (DWT) [5–7] have been applied to transform and analyze brain signals while different mental tasks are performed. Both, EMD and DWT, have shown to be effective in decomposing non-stationary/non-linear time series. But, EMD has the advantage that it does not need the definition of any mother function or pre-processing to improve the signal-to-noise ratio [4]. On the other hand, DWT needs a pre-processing stage to fit the appropriate mother function depending to the task/neuro-paradigm used.

Biometric systems based on EEG-recordings can be separated into states: task-related-state and resting-states. In task-related state different ways have been used to stimulate the brain, for example in [8] Visual Counting and geometric figure Rotation (visual stimulation of images) were used. Another way presented in [9] consisted in mental composition of letters, or as in [10], imagining random digit numbers were presented, among others method and techniques that can be found in [11, 12]. However, persons with certain diseases (e.g., Amyotrophic Lateral Sclerosis, Attention Deficit Disorder, etc) [13, 14] cannot perform some tasks and the use of the above stimuli are not feasible.

The use of EEG signal from resting-states has been reported for example in [7] where a method based on Morlet Wavelet and *Linear SVM* was tested using a dataset of 40 Subjects (192 instances per Subject) and the signal was captured with a sample rate of 256 Hz from 64 channels. In that work, resting-states with the lengths of 300, 60 and 30 seconds were used obtaining accuracies of 1.00, 0.96 and 0.72. However, the use of 300, 60 or even 30 seconds of brain signal length is computationally costly and for a real-time application, fast recognition capabilities with limited information will be essential.

In [15], a method based on Convolutional Neural Networks using the raw signal as input (without pre-processing and without feature extraction) was presented. The dataset was obtained using BCI2000 from 64 channels with a sample rate of 160 Hz. from 10 Subjects and 55 instances of 1 second of duration per subject. Three different experiments were presented: Using resting-states with Open-Eyes/Closed-Eyes/both, and the accuracies obtained were 0.88, 0.86 and 0.82 respectively.

Although the use of resting-states as a biometric marker has been reported by several researchers [7, 11, 15], the possibility of using fewer channels or fewer instances has not been explored so far. As mentioned earlier in the paper, the use of 64 channels does not support the concept of a flexible, low-cost portable EEG-device as presented in [16]. The biometric systems currently adopted by the industry/market use about 5 instances or even fewer to add a new person (e.g., fingerprint, voice/face recognition, retinal scans) and in the research on biometric systems based on EEG, 192 or 55 instances per Subject, which is not practical for a real implementation.

In more recent works, Subject identification methods based on *imagined speech* using DWT [5] and EMD [4] for feature extraction, were presented and the results obtained suggest that EEG of *imagined speech* can be a good candidate as a biometric marker. In the present work, a new conceptual proposition using resting-states (unconstrained rest) in conjunction with fewer EEG channels/instances, is explored. Resting-states can be a valuable biometric marker when the population is large (e.g., in an airport, big enterprises or government organizations) because of its self-reliance and inherent independence from training.

In the following, the new concept of EEG with a reduced number of channels/instances will be presented followed by the proposition of using resting-states (resting-states without restrictions) in conjunction with this new EEG concept as a flexible and affordable portable recognition/authentication system.

2 Towards a low-density EEG Concept: FlexEEG

In order to realize a low-cost and flexible solution for subject identification from brain signals, a new EEG concept is envisioned and presented in this paper. This new EEG concept will be based on a design with a reduced number of channels and the use of wireless dry electrodes to support portability and ease of use. While a laboratory setting and research-grade EEG equipment ensure a controlled environment and high-quality multiple-channel EEG recording, there are applications, situations, and populations for which this is not suitable. Conventional EEG is challenged by high cost (i.e,. computationally costly), high-density, immobility of equipment and the use of inconvenient conductive gels. One consequence of high-density EEG is that interpretation in real-time is not available today. Technological advancements in dry sensor system have opened avenues of possibilities to develop wireless and portable EEG systems with dry electrodes to reduce many of these barriers.

In [16] a new EEG concept of portable (non-invasive) dry single-channel or low-density EEG system, was introduced. While being portable and relying on dry-sensor technology, it will be expected to produce recordings of comparable quality to a research-grade EEG system but with wider scope and capabilities than conventional lab-based EEG equipment. In short, a single more intelligent EEG sensor could defeat high-density EEG. Through this new concept, the range of applications of EEG signals will be expanded from clinical diagnosis and research to health-care, to better understanding of cognitive processes, to learning and education, and to today hidden/unknown properties behind ordinary human activity and ailments (e.g., resting-states, walking, sleeping, complex cognitive activity, chronic pain, insomnia, etc.).

The proposition of real-time Subjects identification using low-density EEG recordings of resting-states will benefit from an EEG device that can offer the flexibility and capabilities envisioned in the FlexEEG concept. The combination of resting-states brain signals with a flexible EEG design with a reduced number of channels will make possible to materialize low-cost and seamless Subject identification within the reach of everyone.

3 Methods

In this section, the methods used with the aim of Subject identification are described in brief. The idea of using resting-states for Subjects identification is motivated by the fact that resting-states are typically used to analyze problems relative to the Subject internal state of mind [13], and this suggests the existence of unique patterns pertaining to the Subject.

According to [17], a stable resting-state (even called resting-state activity) does not necessarily exist, because spontaneous changes in regional neuronal firing occur even when the organism is apparently in rest-state. Also, spontaneous activation can change local blood flow, cause low-frequency blood oxygenation level-dependent signal fluctuations [18]. In other words, the brain is never really at rest [19], and the term only refers to the absence of goal-directed neuronal action with the integration of information of external environment and the Subject internal state, that could be a starting point to discuss why the Subject identification task can work, which in this paper will be done experimentally.

The methods were applied to a dataset of resting-states from the low-cost EMOTIV EPOC device using 14 channels, 8 (P7, P8, O1, O2, F7, F8, T7 and T8), 4 (F7, F8, T7 and T8), 2 (T7 and T8) and 1 channel (T7) that were placed according to the 10-20 international system [20]. Subsets of channels were selected using a *greedy* algorithm [21] as a first attempt to move towards the FlexEEG Concept [16]. This is done in order to analyze the evolution of the accuracy using each time fewer channels, as it is explained later. Additionally, for the set and each subset of channels, experiments were reproduced using 30, 20, 10, 5 and 3 instances per subject, to observe the evolution of accuracy and to

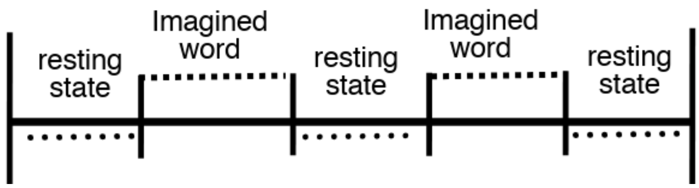

Fig. 1. Protocol for EEG signal acquisition using EMOTIV EPOC [22, 6].

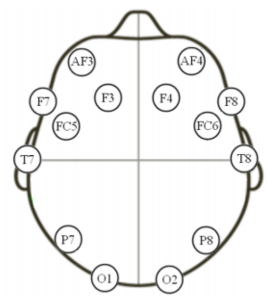

Fig. 2. 10-20 international system for 14 channels

obtain a first approximation of the necessary instances to create a competitive system to the current biometric systems used in industry.

In the following subsections, the methods used are explained in brief, including the dataset description, the logic for channels selection, the feature extraction and classification techniques.

3.1 Dataset description

To test the new conceptual proposition, EEG signals obtained from the low-cost EMOTIV EPOC device with a sample rate of 128 Hz and 14 channels, were used. The dataset consists of brain signals from 27 subjects while imagining 33 repetitions of five imagined words in Spanish, where each repetition was separated by a resting-state. The protocol for acquisition is shown in figure 1 and is described in [22].

The imagined words were recorded in 5 different sessions (not consecutively one after the other), that allows the use of resting-states between instances of words and from 5 different sessions. The mean size of the resting-states in the dataset is $\lceil 3 \rceil$ seconds.

The 14 recorded electrodes as shown in 2 were placed according to the 10-20 international system [20].

Algorithm 1 *Greedy* algorithm for channels selection

procedure CH_SELECTION($subjects$) ▷ 27 Subjects, 14 channels per each one.
  $sj \leftarrow len(subjects)$
  $ch \leftarrow len(sj[0])$
  $ch\_selected \leftarrow [\,[\,]\,]$
  while $ch > 1$ **do** ▷ Stop if there are no channels to remove.
    $ch\_combinations \leftarrow C_{ch,1}$ ▷ k-combinations, $C_{n,k}$
    $accuracies = [\,]$
    for `ch_combination in ch_combinations` **do**
      $accuracies \leftarrow accuracies \cup classifier(subjects, ch\_combination)$
    end for
    $highest\_accuracy \leftarrow max(accuracies)$
    $ch \leftarrow ch\_combinations[highest\_accuracy]$ ▷ $ch \leftarrow ch - 1$
    $ch\_selected \leftarrow ch\_selected \cup ch$
  end while
  return $ch\_selected$ ▷ Channels selected
end procedure

3.2 Channel reduction criteria for low-density EEG

A first step towards the low-density EEG concept discussed in [16] is the channel reduction approach applied based on the information provided for a given neurophysiological task. With this starting point, restrictions of fixed electrodes, the use of a single design for different tasks and its implications for the device portability are discussed. The logic based on the *greedy* algorithm was applied for channels reduction a first step to understand how many channels are needed to obtain sufficient information to detect the relevant activity of the brain.

Channel selection: In the algorithm 1 the *greedy* procedure presented in [21] has been adopted to remove channels in a step-by-step manner. The idea is to obtain the combinations removing 1 channel at a time (k-combinations: $k = 1$) and selecting the subset with the highest accuracy (local maximum). Then the procedure is repeated with the subset obtained while the length of the subset is still greater than 1 channel.

3.3 Feature extraction

Two methods were used for this purpose to compared their capabilities. The first method used is based on the EMD algorithm for which the relevant Intrinsic Mode Functions (IMFs) were decided based on Minkowski distance [23]. Then, for each IMF 4 features were computed: *Instantaneous/Teager energy distribution* and *Higuchi/Petrosian Fractal Dimension*. The flowchart for feature extraction from a given channel is shown in figure 3 and is detailed in [4].

For the second method used, the DWT, first the Common Average Reference (CAR) was applied to improve the signal-to-noise ratio and then the biorthogonal

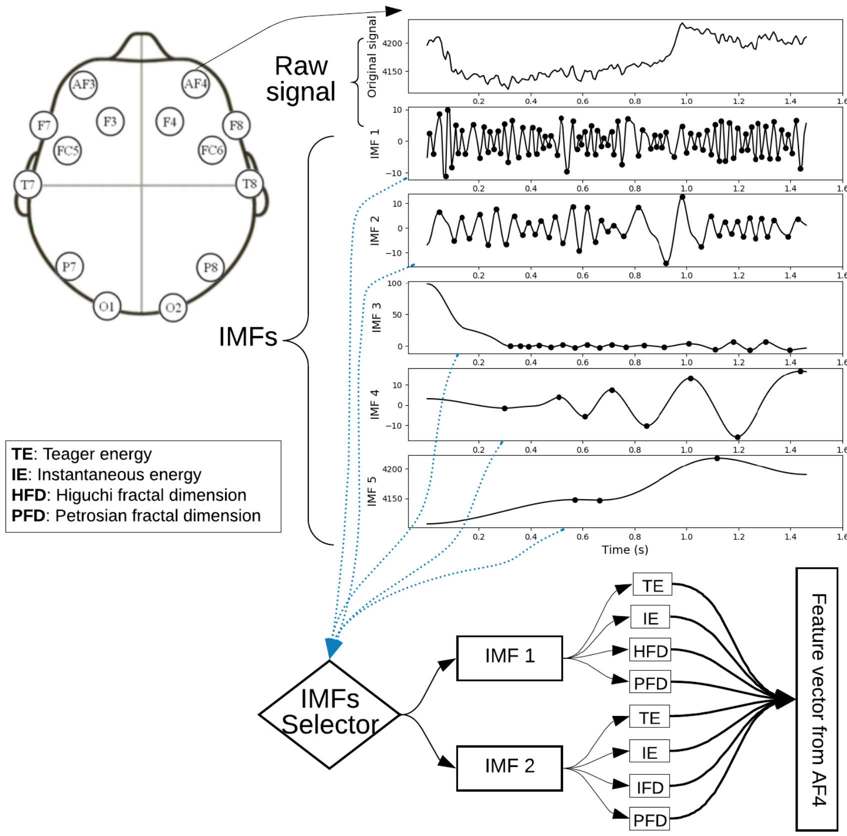

Fig. 3. Flowchart summarizing the feature extraction procedure using EMD

2.2 (bior2.2) DWT with 4 levels of decomposition was computed. Then, for each level of decomposition the *instantaneous energy* was obtained. The flowchart describing the method is shown in figure 4 and is detailed first in [6] and then used for Subject identification using *imagined speech* in [5].

3.4 Classification

The classification procedure was performed using *SVM* (with the kernels: *Linear, Sigmoid and Radial Basis Function*), *Gaussian naive Bayes* and with *k-NN (k=1,2,3,4)*.

To estimate the accuracy and thus evaluate the performance of the methods, {*10, 10, 10, 5, 3*}-folds cross-validation were respectively used for each experiment using respectively 30, 20, 10, 5 and 3 instances per Subject.

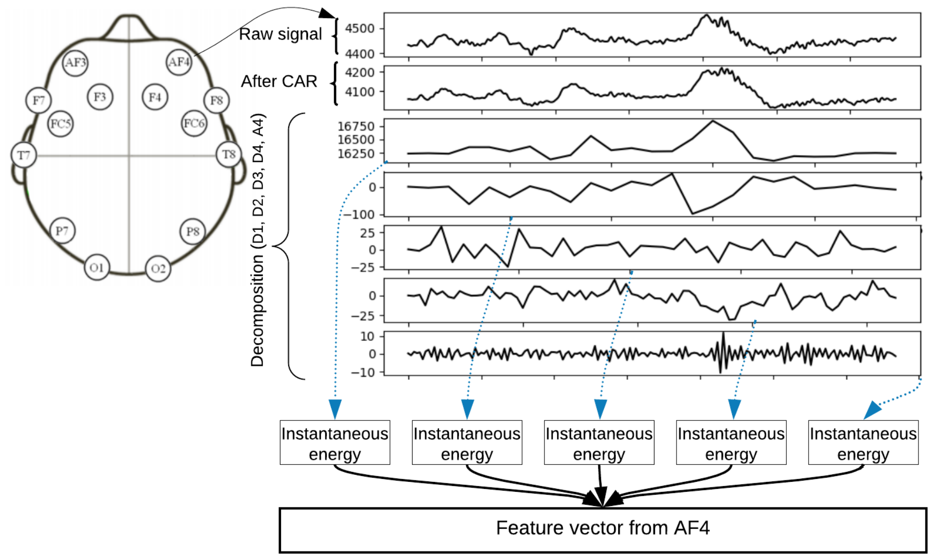

Fig. 4. Flowchart summarizing the feature extraction using bior2.2 DWT

3.5 Experiment setup

The method used to reduce the number of EEG channels provides a general out-work about which channels contain more information and the minimum number of channels for which the loss of accuracy is minimum. However, because the analysis in this paper was carried out between 5 different sessions and using fewer instances; the channels selected in the first sessions/experiment were not the same when using fewer instances or even when using DWT and EMD.

In order to ensure fairness of comparison, the same channels need to be used with EMD and DWT in all sessions and with a different number of instances. Therefore, only the channels that were common to all the experiments were selected. Then, with these subsets of channels, the experiments were repeated with the common channels, to obtain a fair comparison using EMD and DWT.

4 Results

The idea behind the use of fewer channels is to understand and observe the evolution of the accuracy when using DWT and EMD. On the other hand, the method used for channel selection provides a general outlook about which are the channels that contain more information for the Subject identification task and the neuro-paradigm used (In this case: resting-states). According to the method applied for channel selection and after the analysis of the common channels (common channels between instances-used/sessions/methods), the experiments

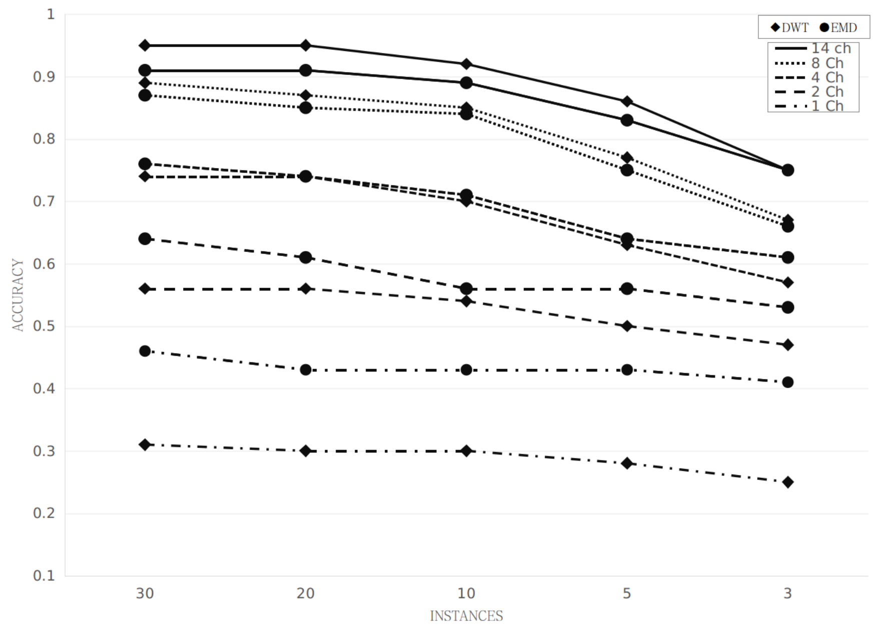

Fig. 5. Average accuracies obtained from 5 sessions using different number of channels with *Linear SVM*.

were repeated using 14 channels, 8 (P7, P8, O1, O2, F7, F8, T7 and T8), 4 (F7, F8, T7 and T8), 2 (T7 and T8) and 1 channel (T7).

In the figure 5 the average accuracies obtained from 5 different sessions with *Linear SVM* and using the set and subsets of channels, are shown.

When 14 and 8 channels were used, the highest accuracies were reached with DWT even when using 3 instances. However using 4, 2 and 1 channels the highest accuracies were reached using EMD. For this task, the evolution of accuracy is clear and easy to understand that EMD can better represent the signal even when the information is reduced. At the same time, these results show that DWT is a more robust method for transforming the brain signals and for feature extraction when the amount of information is higher, which in this context is achievable with a high-density EEG device.

For example, using 14 channels and 30 instances the accuracies obtained with DWT and EMD were 0.95 and 0.91, but using 1 channel and 3 instances the accuracies obtained with DWT and EMD were instead reversed (0.25 and 0.41 respectively) in favor of EMD.

The results presented in figure 6 confirm the observed property above that suggests that the method based on EMD can represent well the brain signal obtaining high accuracy rates.

In figure 7 the results obtained using *Gaussian naive Bayes*, are shown. These results still confirm the high accuracy rates using EMD, but also with an ap-

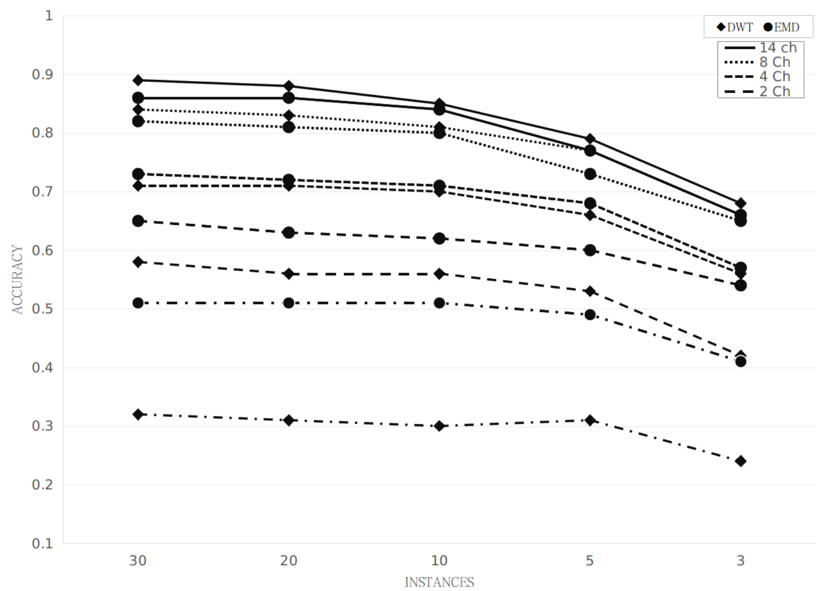

Fig. 6. Average accuracies obtained from 5 sessions using different number of channels with *3-NN*.

parently random behavior when information is low (using 5 and 3 instances per Subject).

A possible response for the *naive Bayes* behavior was presented in [24], where the authors defend the idea that optimal performance is presented in two extreme cases (completely independent features and functionally dependent features) and the performance is worst between these extremes.

In figure 8, the evolution of average accuracies obtained with 20 instances are shown as an example to understand that *SVM* exhibits the highest accuracy using 14 channels but when the number of channels is reduced the *naive Bayes* classifier exhibits higher accuracy rates using EMD compared with the other two classifiers. An important revelation from these result is also that all three classifiers exhibit better accuracies under low-density EEG conditions only when using EMD for feature extraction. In general, using 14 and 8 channels the highest accuracies were obtained using DWT for feature extraction while the highest accuracies using fewer channels (4, 2, 1 channels) were obtained using EMD for feature extraction.

In the tables 1 to 3 the accuracies per session with *Linear SVM, 3-NN (k-NN)* and *Gaussian naive Bayes* using EMD and DWT with a different number of channels and a different number of instances, are presented in detail. The average accuracies (Avg.) and standard deviation (Std) between sessions, also are presented for each subset of channels and using fewer instances.

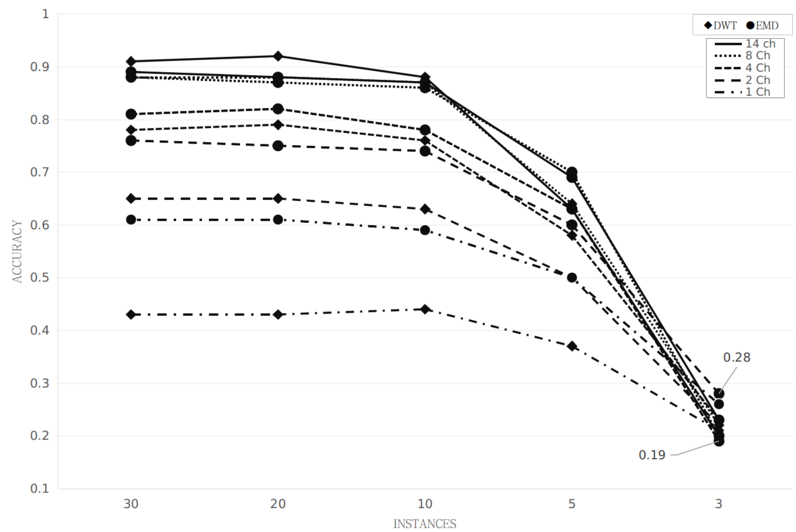

Fig. 7. Average accuracies obtained from 5 sessions using a different number of channels with *Gaussian naive Bayes.*

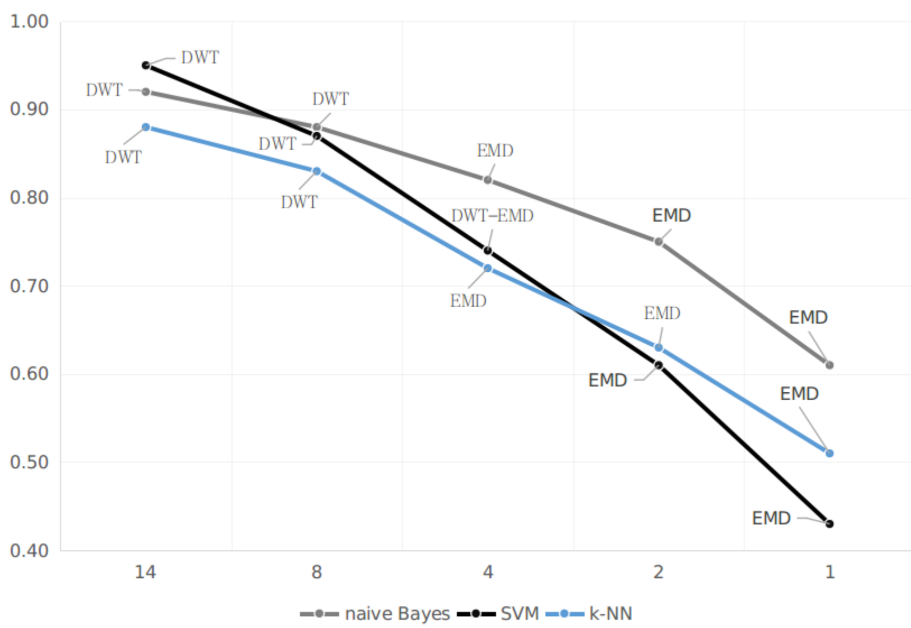

Fig. 8. Evolution of average accuracy using 20 instances with *Linear SVM, k-NN* and *Gaussian naive Bayes.*

Table 1. Accuracies obtained with *Linear SVM* in 5 different sessions using EMD and DWT.

		EMD						DWT					
Session	Ch	30	20	10	5	3	Avg.	30	20	10	5	3	Avg.
1	14	0.89	0.92	0.87	0.83	0.79	0.86±0.05	0.94	0.95	0.94	0.88	0.76	0.89±0.08
	8	0.86	0.87	0.86	0.75	0.69	0.81±0.08	0.90	0.90	0.87	0.84	0.71	0.84±0.08
	4	0.77	0.78	0.79	0.69	0.69	0.74±0.05	0.77	0.79	0.76	0.68	0.68	0.74±0.05
	2	0.65	0.65	0.59	0.62	0.61	0.62±0.03	0.57	0.59	0.58	0.53	0.47	0.55±0.05
	1	0.47	0.48	0.46	0.50	0.47	0.47±0.01	0.32	0.34	0.33	0.26	0.20	0.29±0.06
2	14	0.91	0.91	0.88	0.79	0.81	0.86±0.05	0.95	0.96	0.92	0.87	0.80	0.90±0.07
	8	0.87	0.84	0.84	0.77	0.77	0.82±0.05	0.89	0.88	0.85	0.79	0.71	0.82±0.08
	4	0.76	0.73	0.68	0.63	0.57	0.68±0.07	0.76	0.74	0.72	0.71	0.60	0.71±0.06
	2	0.63	0.57	0.59	0.66	0.52	0.59±0.05	0.61	0.60	0.60	0.54	0.53	0.58±0.04
	1	0.44	0.41	0.42	0.50	0.45	0.44±0.04	0.30	0.28	0.33	0.28	0.33	0.30±0.03
3	14	0.93	0.92	0.91	0.84	0.81	0.88±0.05	0.96	0.96	0.93	0.86	0.75	0.89±0.09
	8	0.89	0.87	0.84	0.80	0.69	0.82±0.08	0.89	0.86	0.85	0.74	0.64	0.80±0.10
	4	0.74	0.71	0.63	0.64	0.57	0.66±0.07	0.70	0.67	0.62	0.58	0.47	0.61±0.09
	2	0.62	0.60	0.52	0.54	0.47	0.55±0.06	0.53	0.51	0.45	0.47	0.39	0.47±0.06
	1	0.48	0.42	0.35	0.37	0.27	0.38±0.08	0.31	0.29	0.26	0.30	0.20	0.27±0.04
4	14	0.92	0.92	0.90	0.86	0.67	0.85±0.11	0.94	0.93	0.88	0.82	0.65	0.85±0.12
	8	0.84	0.84	0.84	0.74	0.59	0.77±0.11	0.86	0.83	0.81	0.72	0.59	0.76±0.11
	4	0.74	0.70	0.66	0.66	0.60	0.67±0.05	0.70	0.72	0.66	0.60	0.53	0.64±0.08
	2	0.60	0.61	0.57	0.54	0.51	0.57±0.04	0.54	0.56	0.55	0.51	0.47	0.52±0.04
	1	0.43	0.39	0.49	0.46	0.33	0.42±0.06	0.28	0.25	0.26	0.22	0.19	0.24±0.04
5	14	0.93	0.89	0.87	0.82	0.69	0.84±0.09	0.95	0.93	0.94	0.86	0.76	0.89±0.08
	8	0.88	0.86	0.83	0.69	0.53	0.76±0.15	0.90	0.90	0.89	0.76	0.69	0.83±0.10
	4	0.80	0.79	0.76	0.61	0.60	0.71±0.10	0.78	0.75	0.72	0.58	0.59	0.69±0.09
	2	0.68	0.63	0.53	0.49	0.56	0.58±0.08	0.56	0.56	0.50	0.46	0.52	0.52±0.04
	1	0.48	0.45	0.42	0.38	0.51	0.45±0.05	0.35	0.33	0.33	0.31	0.31	0.33±0.02
Avg.	14	0.91	0.91	0.89	0.83	0.75		0.95	0.95	0.92	0.86	0.75	
	8	0.87	0.85	0.84	0.75	0.66		0.89	0.87	0.85	0.77	0.67	
	4	0.76	0.74	0.71	0.64	0.61		0.74	0.74	0.70	0.63	0.57	
	2	0.64	0.61	0.56	0.57	0.53		0.56	0.56	0.54	0.50	0.47	
	1	0.46	0.43	0.43	0.44	0.41		0.31	0.30	0.30	0.28	0.25	
Std	14	±0.02	±0.01	±0.02	±0.02	±0.07		±0.01	±0.01	±0.02	±0.02	±0.05	
	8	±0.02	±0.01	±0.01	±0.04	±0.10		±0.02	±0.03	±0.03	±0.05	±0.05	
	4	±0.02	±0.04	±0.07	±0.03	±0.05		±0.04	±0.04	±0.05	±0.06	±0.08	
	2	±0.03	±0.03	±0.03	±0.07	±0.06		±0.03	±0.04	±0.06	±0.03	±0.06	
	1	±0.02	±0.03	±0.05	±0.06	±0.10		±0.03	±0.04	±0.04	±0.03	±0.07	

Inspecting the results in the tables, the following obvious questions come to mind: *What does exactly mean the fluctuation of accuracies? Why sometimes is the accuracy highest with fewer channels/instances?*. An important insight from this paper is that more data is not necessarily more information, and that irrelevant data from certain channels can affect the performance of classifiers depending on the chosen task. Hence the channels selection approach depending on task takes relevance.

4.1 Discussion

The experiments presented were carried out with fewer channels and instances. However, to create a real machine-learning-based model will be necessary to select the best instances to use in a real application. To select those instances the greedy algorithm can be used, but depending on the task and the Subjects, the instances can differ. This means that the selection of instances is Subject-dependent.

Table 2. Accuracies obtained with *3-NN* in 5 different sessions using EMD and DWT.

		EMD						DWT					
Session	Ch	30	20	10	5	3	Avg.	30	20	10	5	3	Avg.
1	14	0.86	0.87	0.88	0.80	0.64	0.81±0.10	0.89	0.89	0.88	0.83	0.76	0.85±0.06
	8	0.84	0.84	0.84	0.79	0.63	0.79±0.09	0.85	0.85	0.84	0.82	0.75	0.82±0.04
	4	0.79	0.79	0.79	0.75	0.64	0.75±0.06	0.75	0.75	0.76	0.73	0.68	0.73±0.03
	2	0.68	0.68	0.67	0.66	0.60	0.66±0.03	0.60	0.57	0.60	0.56	0.40	0.55±0.08
	1	0.51	0.51	0.53	0.55	0.48	0.52±0.03	0.36	0.35	0.33	0.29	0.23	0.31±0.06
2	14	0.85	0.85	0.83	0.72	0.67	0.78±0.08	0.90	0.90	0.84	0.81	0.72	0.83±0.07
	8	0.82	0.81	0.81	0.74	0.72	0.78±0.05	0.84	0.83	0.80	0.81	0.69	0.80±0.06
	4	0.72	0.69	0.67	0.60	0.48	0.63±0.10	0.74	0.73	0.71	0.70	0.60	0.70±0.08
	2	0.65	0.62	0.63	0.62	0.44	0.59±0.08	0.62	0.61	0.61	0.60	0.44	0.57±0.08
	1	0.48	0.49	0.48	0.48	0.40	0.46±0.04	0.34	0.33	0.36	0.42	0.25	0.34±0.06
3	14	0.87	0.88	0.83	0.78	0.72	0.81±0.07	0.89	0.88	0.86	0.80	0.65	0.81±0.10
	8	0.82	0.82	0.77	0.73	0.69	0.77±0.06	0.84	0.83	0.80	0.74	0.55	0.75±0.12
	4	0.71	0.72	0.69	0.69	0.53	0.67±0.08	0.66	0.65	0.63	0.62	0.45	0.60±0.09
	2	0.63	0.59	0.58	0.55	0.53	0.58±0.04	0.52	0.52	0.48	0.50	0.37	0.48±0.06
	1	0.52	0.51	0.44	0.40	0.32	0.44±0.08	0.29	0.27	0.22	0.30	0.20	0.26±0.05
4	14	0.87	0.85	0.82	0.78	0.63	0.79±0.10	0.86	0.83	0.80	0.73	0.55	0.75±0.13
	8	0.79	0.76	0.76	0.70	0.56	0.71±0.09	0.80	0.79	0.78	0.71	0.55	0.73±0.11
	4	0.70	0.68	0.64	0.67	0.57	0.65±0.05	0.69	0.70	0.68	0.58	0.45	0.62±0.11
	2	0.62	0.60	0.58	0.58	0.51	0.58±0.04	0.59	0.57	0.58	0.51	0.36	0.52±0.09
	1	0.49	0.49	0.55	0.52	0.41	0.44±0.05	0.26	0.26	0.22	0.25	0.21	0.24±0.02
5	14	0.87	0.85	0.85	0.76	0.67	0.80±0.09	0.89	0.89	0.85	0.78	0.73	0.83±0.07
	8	0.83	0.81	0.82	0.71	0.63	0.76±0.09	0.85	0.85	0.84	0.80	0.71	0.81±0.06
	4	0.75	0.74	0.76	0.69	0.63	0.71±0.06	0.72	0.72	0.74	0.69	0.63	0.70±0.04
	2	0.68	0.66	0.62	0.58	0.60	0.63±0.04	0.58	0.56	0.54	0.47	0.51	0.53±0.04
	1	0.53	0.55	0.54	0.49	0.44	0.51±0.05	0.35	0.36	0.36	0.30	0.31	0.34±0.03
Avg.	14	0.86	0.86	0.84	0.77	0.66		0.89	0.88	0.85	0.79	0.68	
	8	0.82	0.81	0.80	0.73	0.65		0.84	0.83	0.81	0.78	0.65	
	4	0.73	0.72	0.71	0.68	0.57		0.71	0.71	0.70	0.66	0.56	
	2	0.65	0.63	0.62	0.60	0.54		0.58	0.56	0.56	0.53	0.42	
	1	0.51	0.51	0.51	0.49	0.41		0.32	0.31	0.30	0.31	0.24	
Std	14	±0.01	±0.01	±0.02	±0.03	±0.04		±0.01	±0.03	±0.03	±0.04	±0.09	
	8	±0.02	±0.03	±0.03	±0.03	±0.06		±0.02	±0.02	±0.03	±0.05	±0.09	
	4	±0.04	±0.04	±0.06	±0.05	±0.07		±0.04	±0.04	±0.05	±0.06	±0.10	
	2	±0.03	±0.04	±0.04	±0.04	±0.07		±0.04	±0.03	±0.05	±0.05	±0.06	
	1	±0.02	±0.02	±0.05	±0.06	±0.06		±0.04	±0.05	±0.07	±0.06	±0.04	

In addition, the findings in the experiments suggest that EMD-based method can be used for feature extraction. However, for a certain task, it will be necessary to fix some problems related to the EMD method itself: the well-known mode-mixing problem and explore and test possible solutions available in the state-of-the-art to improve the methodology presented in this paper [26, 27].

5 Conclusions

In this paper, a comparison of EMD and DWT for Subject identification using EEG recordings of resting-states, was presented. In addition, the new FlexEEG Concept [16] was introduced and two of its main challenges were tackled: the use of fewer EEG channels and fewer instances to obtain a competitive/unhackable biometric system. In this paper and as a first attempt to reduce the number of channels, the *greedy* algorithm, was tested. The results obtained are promising and show that resting-states can be effectively used as a biometric marker for subject identification. The experiments conducted on a dataset obtained with

Table 3. Accuracies obtained with *Gaussian naive Bayes* in 5 different sessions using EMD and DWT.

		EMD						DWT					
Session	Ch	30	20	10	5	3	Avg.	30	20	10	5	3	Avg.
1	14	0.88	0.89	0.90	0.78	0.23	0.74±0.29	0.92	0.93	0.90	0.63	0.15	0.71±0.34
	8	0.87	0.88	0.91	0.80	0.20	0.73±0.30	0.89	0.90	0.90	0.62	0.20	0.70±0.31
	4	0.82	0.84	0.85	0.68	0.24	0.69±0.26	0.84	0.85	0.82	0.55	0.15	0.64±0.30
	2	0.77	0.78	0.80	0.65	0.28	0.66±0.22	0.70	0.71	0.73	0.59	0.25	0.60±0.20
	1	0.61	0.63	0.64	0.54	0.29	0.54±0.14	0.47	0.49	0.51	0.41	0.23	0.42±0.11
2	14	0.87	0.85	0.84	0.70	0.32	0.72±0.23	0.92	0.93	0.90	0.64	0.31	0.74±0.27
	8	0.87	0.85	0.87	0.72	0.16	0.69±0.31	0.90	0.89	0.89	0.66	0.28	0.72±0.27
	4	0.79	0.81	0.75	0.63	0.16	0.63±0.27	0.79	0.82	0.80	0.61	0.32	0.67±0.21
	2	0.75	0.74	0.72	0.62	0.25	0.62±0.21	0.70	0.70	0.70	0.57	0.25	0.58±0.19
	1	0.60	0.57	0.58	0.52	0.25	0.51±0.14	0.42	0.43	0.46	0.45	0.21	0.39±0.10
3	14	0.92	0.89	0.87	0.73	0.24	0.73±0.28	0.93	0.94	0.91	0.61	0.25	0.73±0.30
	8	0.89	0.87	0.83	0.70	0.23	0.70±0.28	0.89	0.90	0.87	0.62	0.27	0.71±0.27
	4	0.82	0.80	0.76	0.64	0.17	0.64±0.27	0.75	0.74	0.69	0.56	0.25	0.60±0.21
	2	0.75	0.72	0.68	0.60	0.27	0.60±0.20	0.60	0.59	0.53	0.49	0.17	0.48±0.18
	1	0.60	0.59	0.51	0.46	0.21	0.47±0.16	0.42	0.37	0.36	0.33	0.19	0.33±0.09
4	14	0.90	0.89	0.88	0.69	0.19	0.71±0.30	0.90	0.89	0.86	0.66	0.19	0.70±0.30
	8	0.89	0.88	0.84	0.68	0.17	0.69±0.30	0.86	0.86	0.85	0.62	0.19	0.68±0.29
	4	0.81	0.81	0.75	0.62	0.21	0.64±0.25	0.75	0.75	0.74	0.54	0.20	0.60±0.24
	2	0.75	0.74	0.73	0.58	0.31	0.62±0.19	0.62	0.61	0.59	0.42	0.21	0.49±0.17
	1	0.60	0.61	0.60	0.46	0.25	0.51±0.15	0.39	0.41	0.39	0.31	0.16	0.33±0.10
5	14	0.89	0.88	0.84	0.54	0.20	0.67±0.30	0.88	0.90	0.85	0.62	0.11	0.67±0.34
	8	0.88	0.89	0.85	0.58	0.23	0.69±0.29	0.85	0.87	0.85	0.67	0.15	0.68±0.31
	4	0.83	0.83	0.80	0.58	0.19	0.64±0.28	0.76	0.76	0.74	0.62	0.21	0.62±0.23
	2	0.78	0.77	0.76	0.57	0.28	0.63±0.21	0.62	0.63	0.60	0.44	0.20	0.50±0.18
	1	0.63	0.65	0.62	0.51	0.31	0.54±0.14	0.45	0.46	0.46	0.38	0.27	0.40±0.08
Avg.	14	0.89	0.88	0.87	0.69	0.23		0.91	0.92	0.88	0.63	0.20	
	8	0.88	0.87	0.86	0.70	0.20		0.88	0.88	0.87	0.64	0.22	
	4	0.81	0.82	0.78	0.63	0.19		0.78	0.79	0.76	0.58	0.23	
	2	0.76	0.75	0.74	0.60	0.28		0.65	0.65	0.63	0.50	0.22	
	1	0.61	0.61	0.59	0.50	0.26		0.43	0.43	0.44	0.37	0.21	
Std	14	±0.02	±0.02	±0.02	±0.09	±0.05		±0.02	±0.02	±0.03	±0.02	±0.08	
	8	±0.01	±0.01	±0.03	±0.08	±0.03		±0.02	±0.02	±0.02	±0.02	±0.06	
	4	±0.01	±0.02	±0.04	±0.04	±0.03		±0.04	±0.05	±0.05	±0.04	±0.06	
	2	±0.02	±0.02	±0.04	±0.03	±0.02		±0.05	±0.05	±0.08	±0.08	±0.03	
	1	±0.02	±0.03	±0.05	±0.04	±0.04		±0.03	±0.04	±0.06	±0.06	±0.04	

an EMOTIV EPOC EEG device were reproduced by gradually reducing the number of channels and instances, in the first attempt towards a low-cost EEG-based subject identification. The relatively high accuracies obtained with fewer channels indicate a promising potential towards the FlexEEG concept.

When the methods based on EMD and DWT were compared, the difference in performance increases when using less information. EMD shows a robust and powerful property for feature extraction especially when fewer instances and fewer channels are used. This finding suggests that combining brain signals during resting-states with the use of EMD and the *Gaussian naive Bayes* classifier for low-density EEG, can materialize in a valuable biometric system based on a low-cost EEG-device.

A limitation of the method is that the use of the proposed method will require to find the smallest number of channels and instances to obtain similar accuracies as with high-density EEG. In general, a drastic fall of accuracy is observed from 10 to 5 instances and from 2 (or even 4) to 1 channel. In the future, alternative approaches will be tested for channel selection to improve the performance obtained in these experiments.

Since the focus of this work is the use of EEG-recordings in real-time applications, it is necessary to analyze the computational complexity of the algorithms used to process a signal of size (N): In the case of DWT is $\mathcal{O}(N \log_2 N)$ [25], and for EMD $\mathcal{O}(N \log N)$ [28]. Recently, the real-time implementation of EMD has been reported by some authors [29, 30], and among the challenges anticipated, techniques to distribute the computation and memory, and considerations about the benefits of cloud computing [5, 31] have been discussed.

Future efforts will be directed towards the use of Multivariate Empirical Mode Decomposition [32] which is aimed at multichannel data analysis, but can also be explored for channel selection taking into account the findings of this work.

Acknowledgments. This work was supported by Enabling Technologies - NTNU, under the project "David versus Goliath: single-channel EEG unravels its power through adaptive signal analysis - FlexEEG".

References

1. Jain, Anil K., Arun Ross, and Salil Prabhakar: An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology 14, no. 1 (2004): 4-20.
2. Jain, Anil K., Arun Ross, and Umut Uludag: Biometric template security: Challenges and solutions. Signal Processing Conference 13th European IEEE (2005): 1-4.
3. Valizadeh, Seyed Abolfazl, Franziskus Liem, Susan Mérillat, Jürgen Hänggi, and Lutz Jäncke: Identification of individual subjects on the basis of their brain anatomical features. Scientific reports 8, no. 1 (2018): 5611.
4. Moctezuma, Luis Alfredo and Marta Molinas: EEG-based Subjects Identification based on Biometrics of Imagined Speech using EMD. The 11th International Conference on Brain Informatics (BI 2018) (2018):.
5. Moctezuma, Luis Alfredo, Marta Molinas, A. A. Torres García, Luis Villaseñor Pineda and Maya Carrillo: Towards an API for EEG-Based Imagined Speech classification. International conference on Time Series and Forecasting (2018):.
6. Moctezuma, Luis Alfredo: Distinción de estados de actividad e inactividad lingüística para interfaces cerebro computadora. Thesis project of Master Degree (2017):.
7. Nishimoto, Takashi, Yoshiki Azuma, Hiroshi Morioka, and Shin Ishii: Individual Identification by Resting-State EEG Using Common Dictionary Learning. In International Conference on Artificial Neural Networks, Springer, Cham, (2017): 199-207.
8. Ashby, Corey, Amit Bhatia, Francesco Tenore, and Jacob Vogelstein: Low-cost electroencephalogram (EEG) based authentication. In Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on (2011): 442-445
9. Palaniappan, Ramaswamy: Electroencephalogram signals from imagined activities: A novel biometric identifier for a small population. In International Conference on Intelligent Data Engineering and Automated Learning, Springer, Berlin, Heidelberg (2006): 604-611.

10. Jayarathne, Isuru, Michael Cohen, and Senaka Amarakeerthi: BrainID: Development of an EEG-based biometric authentication system. In Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016 IEEE 7th Annual (2016): 1-6.
11. Jayarathne, Isuru, Michael Cohen, and Senaka Amarakeerthi: Survey of EEG-based biometric authentication. In Awareness Science and Technology (iCAST), 2017 IEEE 8th International Conference on (2017): 324-329.
12. Del Pozo-Banos, Marcos, Jesús B. Alonso, Jaime R. Ticay-Rivas, and Carlos M. Travieso: Electroencephalogram subject identification: A review. Expert Systems with Applications 41, no. 15 (2014): 6537-6554.
13. Elman, Lauren B., and L. McCluskey: Clinical features of amyotrophic lateral sclerosis and other forms of motor neuron disease. Up-to-date. Waltham: Wolters Kluwer Health (2012): 23.
14. Feller, T. G., R. E. Jones, and M. G. Netsky: Amyotrophic lateral sclerosis and sensory changes. Virginia medical monthly **93** , no. 6 (1966): 328.
15. Ma, Lan, James W. Minett, Thierry Blu, and William SY Wang: Resting State EEG-based biometrics for individual identification using convolutional neural networks. In Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE (2015): 2848-2851.
16. Molinas, Marta, Audrey Van der Meer, Nils Kristian Skjærvold and Lars Lundheim: David versus Goliath: single-channel EEG unravels its power through adaptive signal analysis - FlexEEG. Research project (2018):.
17. Xiong, Jinhu, Liangsuo Ma, Binquan Wang, Shalini Narayana, Eugene P. Duff, Gary F. Egan, and Peter T. Fox: Long-term motor training induced changes in regional cerebral blood flow in both task and resting states. Neuroimage 45, no. 1 (2009): 75-82.
18. Golanov, EUGENE V., S. E. I. J. I. Yamamoto, and DONALD J. Reis: Spontaneous waves of cerebral blood flow associated with a pattern of electrocortical activity. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 266, no. 1 (1994): R204-R214.
19. Mantini, Dante, Mauro G. Perrucci, Cosimo Del Gratta, Gian L. Romani, and Maurizio Corbetta: Electrophysiological signatures of resting state networks in the human brain. Proceedings of the National Academy of Sciences 104, no. 32 (2007): 13170-13175.
20. Jasper, Herbert: Report of the committee on methods of clinical examination in electroencephalography. Electroencephalogr Clin Neurophysiol **10** (1958): 370-375.
21. Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein: Introduction to algorithms, chapter 16 “Greedy Algorithms”. MIT press, (2001):.
22. Torres-García, A. A., C. A. Reyes-García, L. Villaseñor-Pineda, and J. M. Ramírez-Cortís: Análisis de señales electroencefalográficas para la clasificacin de habla imaginada. Revista mexicana de ingeniería biomédica **34**, no. 1 (2013): 23-39.
23. Boutana, Daoud, Messaoud Benidir, and Braham Barkat: On the selection of intrinsic mode function in EMD method: application on heart sound signal. In Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on (2010): 1-5.
24. Rish, Irina, Joseph Hellerstein, and Jayram Thathachar: An analysis of data characteristics that affect naive Bayes performance. IBM TJ Watson Research Center 30 (2001):.
25. Averbuch, Amir Z., and Valery A. Zheludev: Construction of biorthogonal discrete wavelet transforms using interpolatory splines. Applied and Computational Harmonic Analysis **12**, no. 1 (2002): 25-56.

26. Gao, Yunchao, Guangtao Ge, Zhengyan Sheng, and Enfang Sang: Analysis and solution to the mode mixing phenomenon in EMD. In Image and Signal Processing, 2008. CISP'08. Congress on, vol. 5 (2008): 223-227.
27. Fosso, Olav B., and Marta Molinas: Method for Mode Mixing Separation in Empirical Mode Decomposition. arXiv preprint arXiv:1709.05547 (2017).
28. Wang, Yung-Hung, Chien-Hung Yeh, Hsu-Wen Vincent Young, Kun Hu, and Men-Tzung Lo: On the computational complexity of the empirical mode decomposition algorithm. Physica A: Statistical Mechanics and its Applications **400** (2014): 159-167.
29. Fontugne, Romain, Pierre Borgnat, and Patrick Flandrin: Online empirical mode decomposition. In Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on (2017): 4306-4310.
30. Faltermeier, Rupert, Angela Zeiler, Ingo R. Keck, Ana Maria Tom, Alexander Brawanski, and Elmar Wolfgang Lang: Sliding empirical mode decomposition. In Neural Networks (IJCNN), The 2010 International Joint Conference on. (2010): 1-8.
31. Mahmudova, Shafagat: Analysis of biometric authentication methods of users in clouds. International Journal of Advanced Engineering and Technology, **1**, no. 5 (2017): 14-17.
32. Rehman, Naveed, and Danilo P. Mandic: Multivariate empirical mode decomposition. In Proceedings of The Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 466, no. 2117 (2010): 1291-1302.

View publication stats