UNIVERSITY OF WEST LONDON
The Career University

UWL REPOSITORY

repository.uwl.ac.uk

Task sensitivity in EEG biometric recognition

Yang, Su ORCID logoORCID: https://orcid.org/0000-0002-6618-7483, Deravi, Farzin and Hoque, Sanaul (2018) Task sensitivity in EEG biometric recognition. Pattern Analysis and Applications, 21 (1). pp. 105-117. ISSN 1433-7541

http://dx.doi.org/10.1007/s10044-016-0569-4

This is the Published Version of the final output.

UWL repository link: https://repository.uwl.ac.uk/id/eprint/7926/

Alternative formats: If you require this document in an alternative format, please contact: open.research@uwl.ac.uk

Copyright: Creative Commons: Attribution 4.0

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy: If you believe that this document breaches copyright, please contact us at open.research@uwl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Rights Retention Statement:

Pattern Anal Applic (2018) 21:105–117
https://doi.org/10.1007/s10044-016-0569-4

THEORETICAL ADVANCES

Task sensitivity in EEG biometric recognition

Su Yang[1] · **Farzin Deravi**[1] · **Sanaul Hoque**[1]

Received: 8 December 2015 / Accepted: 11 July 2016 / Published online: 29 July 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract This work explores the sensitivity of electroencephalographic-based biometric recognition to the type of tasks required by subjects to perform while their brain activity is being recorded. A novel wavelet-based feature is used to extract identity information from a database of 109 subjects who performed four different motor movement/imagery tasks while their data were recorded. Training and test of the system was performed using a number of experimental protocols to establish if training with one type of task and tested with another would significantly affect the recognition performance. Also, experiments were conducted to evaluate the performance when a mixture of data from different tasks was used for training. The results suggest that performance is not significantly affected when there is a mismatch between training and test tasks. Furthermore, as the amount of data used for training is increased using a combination of data from several tasks, the performance can be improved. These results indicate that a more flexible approach may be incorporated in data collection for EEG-based biometric systems which could facilitate their deployment and improved performance.

Keywords EEG · Biometrics · Identification · Verification

✉ Sanaul Hoque
S.Hoque@kent.ac.uk

Su Yang
sy91@kent.ac.uk

Farzin Deravi
F.Deravi@kent.ac.uk

[1] School of Engineering and Digital Arts, University of Kent, Canterbury, UK

1 Introduction

With the rapid development of machine learning techniques as well as the increasing availability of low-cost sensors, the biometric person recognition technologies have become an active area of research in recent years, leading to significant deployments in a range of application domains. However, despite some considerable successes, important challenges still hinder their widespread adoption and acceptance [1], and because of this the search for new biometric modalities continues. Bio-signals are potentially rich in identity information, which make them appealing candidates for biometric applications. With its nonstationary characteristics [2], the electroencephalographic (EEG) signal is becoming an attractive choice as a biometric modality in some applications due to its natural resistance to spoofing and increasing ease of acquisition through low-cost sensors.

While EEG-based biometrics would require the use of dedicated sensors and a measure of cooperation from the users to ensure consistent and reliable signals that can be captured to recognize their identity, there are special use-cases where this modality could be of practical use. One can consider scenarios where the use of a headset is natural in user interactions, such as driving certain motor vehicles, or during the performance of certain activities (e.g., safety/security tasks) where the sensors could be built in the helmet/headset that the user will naturally wear. In this case, verification of identity can take place in a hands-free manner as and when required. Use of this modality can also ensure a greater degree of counter-spoofing through continuous liveness detection. The literature survey that follows presents a picture of active and growing interest in this biometric modality.

Poulos et al. [3–5] first proposed to employ EEG signals for person identification. Since then, this modality has been

increasingly receiving attention in its potential biometric applications. Intuitively, EEG signals are expected to contain some information unique to individuals. However, it is not clear what deliberate or involuntary mental activity would generate the best and most biometrically informative signals. This question is closely related to which scalp region should provide the signals for biometric recognition.

The mental activity or motor movement tasks used in the research literature to trigger EEG signals for biometric processing could generally be grouped into three main categories:

1. Resting state, with no intentional mental or physical activity with eyes either open or closed,
2. Event-Related Potential (ERP) signals, especially the P300 evoked potential [6] triggered by visual stimuli or motor movement, and
3. Intentional mental activity(s), such as mental counting or motor imagery.

Some important research results related to these three categories are reviewed briefly in the following subsections.

1.1 Resting state EEG

Su et al. [7] reported their system's performance while using only the Fp1 electrode position (frontal region) for data collection from 40 healthy subjects, while participants rested on a sofa with their eyes closed. Each subject provided 60 min (12 recordings) of recordings in total, and half of this data was randomly used for training and the rest was used for testing. A correct recognition rate (CRR) of 97.5 % was reported.

Lee et al. [8] also captured their EEG data while subjects were resting with their eyes closed but only four subjects were included. Data were obtained using O1 electrode (occipital region) in two sessions with the time intervals ranging from 10 days to 5 months. Data of the first session were used as the training set, and the second session's recording was used for testing (20 s of training and 20 s of testing). An accuracy of 98.33 % was achieved.

Recently, Rocca et al. [9] reported the EEG identification performance using a relatively large database. Two subsets of a publicly available database of 108 subjects in resting state were analysed, one with eyes open and the other with eyes closed. There was 1-min-long EEG recording for each subject, and ten seconds of test data was used for sixfold leave-one-out cross-validation. A performance of 100 % recognition accuracy was reported using the fusion of conventional power spectral feature and their proposed functional connectivity feature.

A potential problem of using the EEG data captured during the resting state for biometrics recognition may be the ambiguity of the instruction given to the participants during the data collection, which may be interpreted by the subjects in different ways, resulting in incommensurable data.

1.2 Event-stimulated EEG

An event-related potential (ERP) is the measured brain response that is the direct result of a specific sensory, cognitive, or motor event [10]. The P300 wave is one such ERP component obtained during the process of decision-making, such as the reaction to the *oddball paradigm* [11]. In such a visual evoked potential (VEP) setting, the visual stimulus results in an EEG P300 signal. The P300 signal appears as a positive deflection in the measured EEG voltage with latency (delay between stimulus and response) roughly in the range of 250–500 ms [6]. Researchers have used this particular waveform for biometric recognition.

Using the P300 signal directly from a single electrode, Singhal et al. [12] reported an average identification accuracy of 78 % for a database containing 10 subjects. A "peak matching algorithm" was applied to the averaged VEP signal in the time domain for comparison. Yearn et al. [13] also investigated VEP signals generated using a face stimuli in an authentication scenario using a data set containing ten subjects and captured over 2 sessions conducted on different days. An equal error rate of 14.5 % was achieved using 18 electrodes.

Palaniappan et al. [14] employed the P300 VEP for feature extraction, while people were viewing a set of pictures originally proposed in [15]. The experiment comprised of 10 subjects using an EEG cap of 61 electrodes. The maximum identification rate achieved was 95 % for the data recorded in a single session. Similar but improved approaches have been tested using databases with larger populations leading to a performance of 98.12 % with a database of 102 subjects [16–18].

1.3 Mental imagery EEG

Considering the limitations of the resting state and the visual stimulus approaches to EEG stimulation for biometric applications, it is only natural for researchers to explore other approaches which may be more controllable than the resting state scenario and potentially less complex than the visual stimulus scenario. The use of mental activity in an identification scenario was first reported in 2005, when EEG data were recorded (from 4 subjects) during the performance of mental tasks (including mathematics, geometric figure rotation, mental letter composing and visual counting) [19].

Marcel et al. [20] used the data captured while imagining hand movements for a biometric authentication

scenario. Power spectral density (PSD) features of the EEG signal were compared using Gaussian mixture models (GMM). Sixteen minutes of recordings from 8 electrodes were used for training and 4 min for testing, and a half total error rate (HTER) of 7.1 % was reported for 9 subjects.

One drawback of employing EEG as a biometric modality has been the complexity of setting up the data acquisition system, given the number of electrodes involved, the time required for their attachment, and the expensive hardware required. Thus, the use of low-cost sensors becomes an important research trend despite the likely reduction in signal quality. Chuang et al. [21] reported a system which employed only a single Fp1 electrode (NeuroSky MindSet [22]); two 40–50-min data collection sessions were conducted on separate days. Different mental activities were performed, and an EER of 1 % was achieved for a database of 15 subjects. However, the identification accuracy was only 22 % when using the same database.

Template ageing effects when using mental tasks with long time interval between training and test sessions have been reported. In [23], EEG data were recorded with 53 electrodes from 9 subjects in two sessions (with motor task data of imaginary finger movements) with a time interval of approximately 1 year. Using part of the data from the first session for training and the rest of the same session's data as the test set, the CRR reached as high as 98 %; whereas using the first session's data for training and data of the second session for testing, the performance reduced to a CRR of 87.1 %.

One possible drawback of using visual stimulus for biometric applications is the need for an external stimulus to trigger the VEP signals. This may make the resulting biometric system more complex compared with alternatives based on using the resting state or directed mental activity. In contrast, for EEG signals captured during the resting state as well as those obtained during the performance of mental/cognitive tasks there is the problem of the variability associated with the users' interpretations of the instruction given.

An extensive review of EEG signals used for biometric recognition can be found in [44]. The published research using mental tasks for generating biometric EEG signals has not considered the impact of task types on performance. The impact of the type of task on the performance that can be achieved in biometric recognition may be significant and is yet to be investigated. Four specific questions are addressed in this work: (1) does the optimal placement of electrodes vary with the movement/imagery task required of the subjects? (2) Does the type of movement/imagery task performed by subjects affect the biometric recognition performance? (3) Would training with data from one task and testing with data from another task significantly affect the performance? (4) Whether combining data from different types of tasks for training of the system improves performance?

The paper is structured as follows: Sect. 2 describes a wavelet-based method for EEG feature extraction and provides details of the particular wavelet features used in this work. Section 3 contains the proposed experimental protocols, which are especially designed to investigate the questions raised in Sect. 1. The experimental evaluations and the analysis of results are included in Sect. 4. Conclusions and suggestions for further work are presented in Sect. 5.

2 EEG biometric system

The block diagram of the proposed EEG-based biometric system used for exploring the task sensitivity is depicted in Fig. 1. Users are instructed to conduct certain motor movement/motor imagery tasks while their EEG data are being recorded. Time and frequency domain features are extracted and subsequently used for user recognition. The performance of the system crucially depends on the choice of features. In this section, we present a wavelet-based technique for feature extraction that will be used in later experiments on task sensitivity.

Similar to certain modalities in the field of signal processing (speech recognition, for example [24]), the EEG signal is also considered nonstationary [2, 25]. Fourier transform (FT) is a conventional approach in signal processing and is widely used for EEG-based signal analysis. However, its use is based on the assumption that the data to be analysed are strictly stationary. Short-time Fourier transform (STFT) may moderately relax this restrictive criterion: by segmenting the nonstationary signal into a series of overlapped short time frames, assuming the data within each frame is stationary and the Fourier transforms is applied to each of these frames separately. This approach, however, may not be able to fully capture the nonstationary dynamics of the signals' content.

In recent decades, the Wavelet Transform (WT) and its related applications have received increasing attention due to its capability of capturing the signal information in both time and frequency domains [26]. By mapping the signal $x(t)$ into a particular space (wavelet space) with a scale a and a shift b, it is possible to reveal both time and frequency content of nonstationary data simultaneously (alleviate the trade-off in FT). The transformation process can be expressed as follows [27]:

$$WT_{\psi}\{x\}(a,b) = \langle x, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} x(t) \cdot \psi_{a,b}(t) \mathrm{d}t \qquad (1)$$

Where $\psi_{a,b}(t)$ is the scaled and shifted version of a given wavelet function:

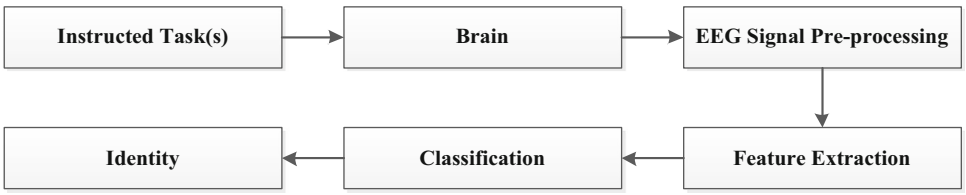

Fig. 1 EEG-based biometric system

Table 1 Wavelet packet decomposition for the proposed system

Decomposition	Wavelet frequency bands							
Level 0	0–80 Hz							
Level 1	0–40 Hz				40 Hz–80 Hz			
Level 2	0–20 Hz		20 Hz–40 Hz		40 Hz–60 Hz		60 Hz–80 Hz	
Level 3	0–10 Hz	10 Hz–20 Hz	20 Hz–30 Hz	30 Hz–40 Hz	40 Hz–50 Hz	50 Hz–60 Hz	60 Hz–70 Hz	70 Hz–80 Hz

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{2}$$

the wavelet coefficients $WT_{\psi}\{x\}(a, b)$ in (1) may theoretically reveal both the time and the frequency properties of signals.

One advantage of the WT is the flexibility of choosing the wavelet functions. Rather than representing the signals by a series of sinusoidal functions, WT decomposes the signal using a series of scaled and shifted wavelet functions; different wavelets may be used based on particular applications [27, 28].

In this work, we propose to employ one discrete form of WT, the wavelet packet decomposition (WPD) transform, which includes a full decomposition of the signals into multiple levels using both wavelet and scaling functions [27]. In conventional WT, each level is calculated by passing only the previous wavelet approximation coefficients through discrete-time low- and high-pass quadrature mirror filters. In the WPD, both the detail and approximation coefficients are decomposed to create the full binary tree [27, 28]. The EEG signals were decomposed up to level 3 (see Table 1). This allows the signal to be divided into eight nonoverlapped wavelet bands. In order to maximize the use of both time and frequency properties of the signal, the coefficients from both level 2 and level 3 were employed as the primary features in this work.

Different decomposition levels result in a series of coefficients with different lengths: the higher the decomposition level, the more frequency details are reflected by the coefficients, hence less time domain information may be retained. Therefore, the coefficients of the four sub-bands from level 2 were retained for feature extraction as well as those of level 3, since they may better retain useful time domain properties of the signal. Based on preliminary investigations, the Daubechies 4 wavelet function was used and a segmentation window size of 4800 samples (30 s) was chosen [29]. Only the bands in Level 2 and Level 3 of Table 1 had been used for feature extraction in this study.

The overall system is illustrated in Fig. 2. The acquired EEG data from I electrodes are segmented in time into N overlapping windows; each window overlaps its neighbour by 50 %. For a given time window, data from each of the electrodes are transformed using multi-level WPD followed by a feature enhancement stage where the derivatives of the WPD coefficients are computed. For each of these feature-enhanced bands, the standard deviation (SD) is calculated. The SDs for all the bands and all the electrodes are then concatenated to produce the feature vector for classification using an LDA classifier. The classifier decisions from all the time windows are fused using the majority voting rule. The performances of this system were investigated for identification and verification scenarios.

Before conducting experiments using this database to explore the sensitivity of the biometric system to task type, it is helpful to verify that there are indeed some significant differences in the four mental/imagery tasks that it includes.

The mean of wavelet coefficients is used as a feature for task discrimination. Data of multiple subjects (first 15 subjects of MM/I data set) were analysed, and the values for the four motor/imagery tasks were plotted. As examples, Fig. 3 depicts the four task clusters for Subject 1 (S1) and Subject 2 (S2) using the first three feature dimensions. It is clear that the clusters of T2 and T4 are close to each other and away from both T1 and T3.

3 Experimental protocols

Data from the "EEG Motor Movement/Imagery Dataset" (MM/I) have been used for the investigations [30, 31]. This data set contains EEG data collected using the BCI 2000

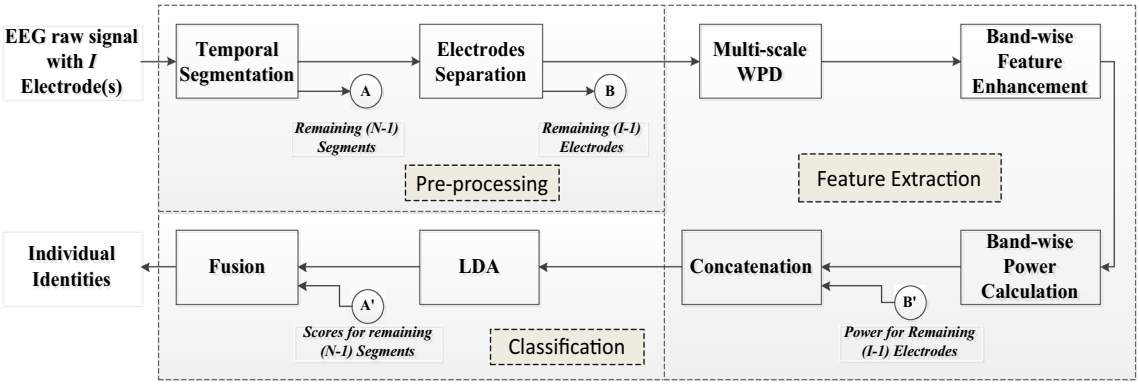

Fig. 2 System Diagram

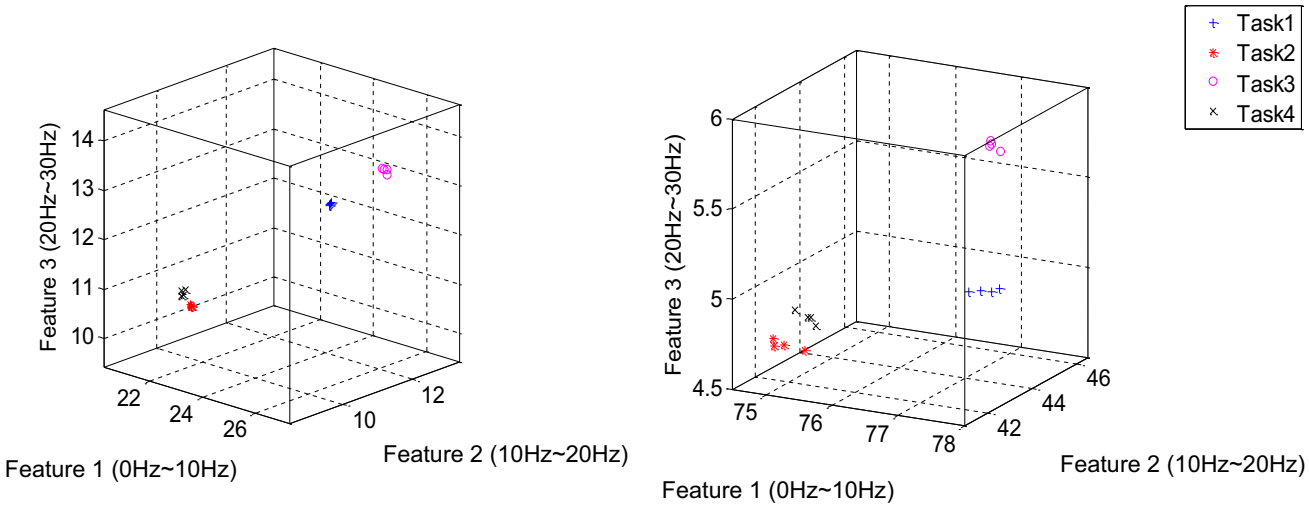

Fig. 3 Clusters of the first three dimensions of the feature vector [from four windows of 30 s duration, S1 (*left*) and S2 (*right*)]

system (sampling frequency 160 Hz) from 109 subjects. In order to guarantee equal and sufficient recording length (at least 2 min for mental tasks), 108 out of 109 subjects' data were selected for the experiments—excluding 1 subjects with shorter data recordings. Subjects performed four different movement/imagery tasks (T1–T4). Additionally two baseline tasks (Tb) were also performed where subjects were in a resting state with both eyes open (EO) and both eyes closed (EC). The four movement/imagery tasks lasted for about 2 min per recording. The motor movement/imagery tasks were repeated three times (three "runs": R1, R2 and R3). The two baseline tasks lasted only 1 min with only a single recording. In brief, the four task instructions given to the subjects to perform are as follows:

- Task 1 (T1)–"open and close left or right fist";
- Task 2 (T2)–"imagine opening and closing left or right fist";
- Task 3 (T3)–"open and close both fists and both feet";
- Task 4 (T4)–"imagine opening and closing both fists and both feet."

Further details of the database may be found in [32].

The first goal of this work is to investigate the biometric performance achieved when using EEG signals from different scalp regions. Nine electrodes clustered in three distinctive scalp regions were selected for analysis (AF3, AFz and AF4 in the frontal lobe (F); C1, Cz and C2 in the motor cortex (M); and O1, Oz and O2 in the occipital lobe (O)). The positioning of the sensors is illustrated in Fig. 4 [33]. These regions were chosen to cover the anatomically significant areas of the brain involved in motor/imagery tasks [34] and to investigate the impact from other regions that are less likely to be activated by the chosen tasks [35]. These abbreviations are combined, using the convention Task-Recording-Region, to generate labels for the data subsets used in the experiments, e.g., TbEOF meaning baseline task with eyes open and data from the frontal region electrodes and T1R1M refers to data from Task 1, Run 1 and motor cortex region electrodes. If a scalp region is not specified in the label, it implies that data from all nine electrodes have been used in the experiment.

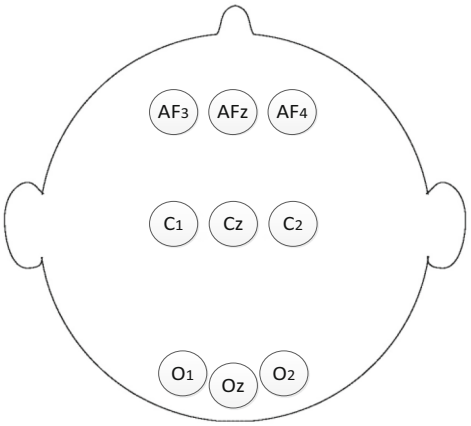

Fig. 4 Chosen electrode positions

Three experimental protocols are proposed to investigate the research questions raised in Sect. 1. These protocols are used for system evaluation in both identification and verification scenarios.

3.1 Protocol P1: region/task pairing

The goal of this protocol (P1) is to investigate the impact of the pairing of electrode regions and task types on system performance. Experiments performed using this protocol will also serve as a preliminary investigation to find the tasks with greatest biometric potential to be investigated further. The training and test data sets for P1 are shown in Table 2. The data subsets identified in P1 make it possible to explore the performance in each electrode region separately. The data from R1 together with R3 are chosen as the training data, and R2's data are employed for testing. P1 also identifies four groups of data subsets matching the four types of motor movement/imagery tasks (T1–T4) in MM/I to facilitate experiments to see the relative performance of each task type.

3.2 Protocol P2: mismatched training/testing tasks

The purpose of the second protocol (P2) is to investigate the impact of using different motor/imagery tasks for training and testing of the system—the test data have been taken from a different task type to that used for training the system (all nine of the selected electrodes are used). P2 makes it possible to see if a mismatch between the training and testing task types can significantly affect the performance.

As in P1, the data from R1 together with R3 are selected for training and R2's data are employed for testing. Additionally, the data from the two baseline data sets are also used in this protocol for testing as illustrated in Fig. 5. In this figure, each arrow signifies a pairing of a training subset and a test subset that is used in experiments. The matched pairings are also included here for comparison.

Table 2 Protocol P1

Experiments	Training set	Test set
P1.1	T1R1F + T1R3F	T1R2F
P1.2	T1R1M + T1R3M	T1R2M
P1.3	T1R1O + T1R3O	T1R2O
P1.4	T1R1 + T1R3	T1R2
P1.5	T2R1F + T2R3F	T2R2F
P1.6	T2R1M + T2R3M	T2R2M
P1.7	T2R1O + T2R3O	T2R2O
P1.8	T2R1 + T2R3	T2R2
P1.9	T3R1F + T3R3F	T3R2F
P1.10	T3R1M + T3R3M	T3R2M
P1.11	T3R1O + T3R3O	T3R2O
P1.12	T3R1 + T3R3	T3R2
P1.13	T4R1F + T4R3F	T4R2F
P1.14	T4R1M + T4R3M	T4R2M
P1.15	T4R1O + T4R3O	T4R2O
P1.16	T4R1 + T4R3	T4R2

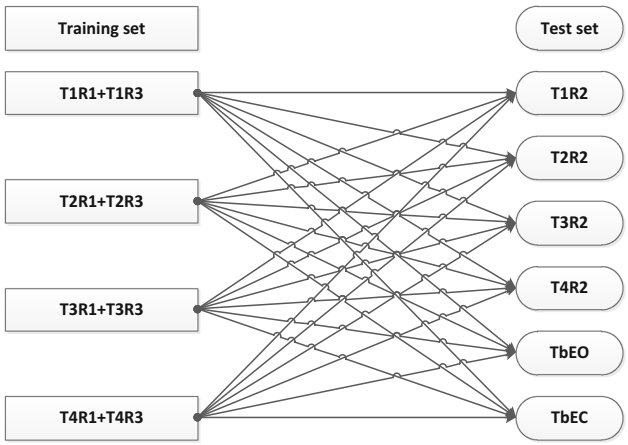

Fig. 5 Protocol P2—Mismatched training/testing tasks

3.3 Protocol P3: heterogeneous training

This protocol (P3) explores if data from different task types may be combined for the training of the system to achieve a better performance. Test data from just one task type and recording (T1R2) was used in this protocol. The training data were generated by including an increasing quantity of data from different task types. The data subsets used in P3 for training and testing are shown Table 3.

Table 3 Protocol P3

Experiments	Training set	Test set
P3.1	T1R1	T1R2
P3.2	T1R3	T1R2
P3.3	T1R1 + T1R3	T1R2
P3.4	T2R1 + T2R2 + T2R3	T1R2
P3.5	T1R1 + T1R3 + T2R1	T1R2
P3.6	T1R1 + T2R1 + T2R2 + T2R3	T1R2
P3.7	T1R1 + T1R3 + T2R1 + T2R2	T1R2
P3.8	T1R1 + T1R3 + T2R1 + T2R2 + T2R3	T1R2
P3.9	T1R1 + T1R3 + T2R1 + T2R2 + T2R3 + T3R1	T1R2
P3.10	T1R1 + T1R3 + T2R1 + T2R2 + T2R3 + T3R1 + T4R1	T1R2
P3.11	T1R1 + T1R3 + T2R1 + T2R2 + T2R3 + T3R1 + T4R1 + T4R2	T1R2

4 Experimental analysis

This section presents and analyses the results from the experiments defined in the protocols in Sect. 3. Both the identification and the verification scenarios are investigated.

4.1 Identification scenario

4.1.1 Test results for P1

Figure 6 presents the performance of the system for different pairings of electrode positions and tasks. The results were generated by randomly selecting 75 % of the EEG data from R1 and R3 to train the system, and the data from R2 were used for testing. The tests were repeated 100 times for generating the box plots.

It is evident that the accuracies achievable from the isolated scalp regions are very comparable for the tasks investigated, although some small variations can be seen. For isolated regions, task T1 produced the highest accuracy (about 89 %) when using data from the Occipital (O) region only. Further inspecting the median accuracies of the four tasks, it appears that introducing feet movements actually adversely affected the biometrics performance: the performance of T3 and T4 (movement or imagery movement of both fist and feet) are both worse than that of T1 and T2 (movement or imagery movement of only fist). When features from all the regions are combined, there is a significant rise of about 7–12 % in the median accuracies for all the tasks. Task T1 in this case had produced the highest median accuracy of about 96 %.

In summary, the results show that EEG data contain adequate discriminatory information to be used for biometrics identification. While the position of the isolated sensors did not make a substantial difference in identification performance, the choice of the task of opening and closing the fists (T1) seems to outperform all the other tasks in the database when all the three regions are used.

4.1.2 Test results from protocol P2

For the tests in Protocol P2, the principal objective was to see the effect of nonmatching training and test tasks on system performance. Here, the features extracted from all the nine electrodes are concatenated for this evaluation. The two baseline resting state recordings (TbEO and TbEC) are also included for testing to establish the usability of such data in conjunction with movement/imagery data used for training in a biometric context. The results are shown in Table 4. It is clear that the performances observed with the baseline resting state tasks used for testing are very poor. This could be due to the fact that training data based on movement/imagery tasks are substantially different in nature from EEG signals obtained in resting state.

On the contrary, when the system was tested with nonmatching movement/imagery tasks, the performances were very promising and comparable to each other. In particular, training with the data from task T1 has again shown the highest identification accuracies amongst the four tasks for nonmatching training/testing scenarios explored here. In some cases, the nonmatching data sets actually showed better accuracy than those from the matching data sets (e.g., training by T1 and test with T2 provided the best performance).

In short, the results suggest that given a particular type of motor movement/imagery task used for preparing the training data, the system may still be able to give acceptable results while tested by a different movement/imagery task data. This allows more flexibility from the perspective of both system designers and users in real-life biometric applications. The impact on recognition performance from different motor movement/imagery tasks data is limited, whereas using the resting state EEG for testing was found to be ineffective.

4.1.3 Test according to P3

The results of the previous experiments have shown that the match between the training and the testing task types is

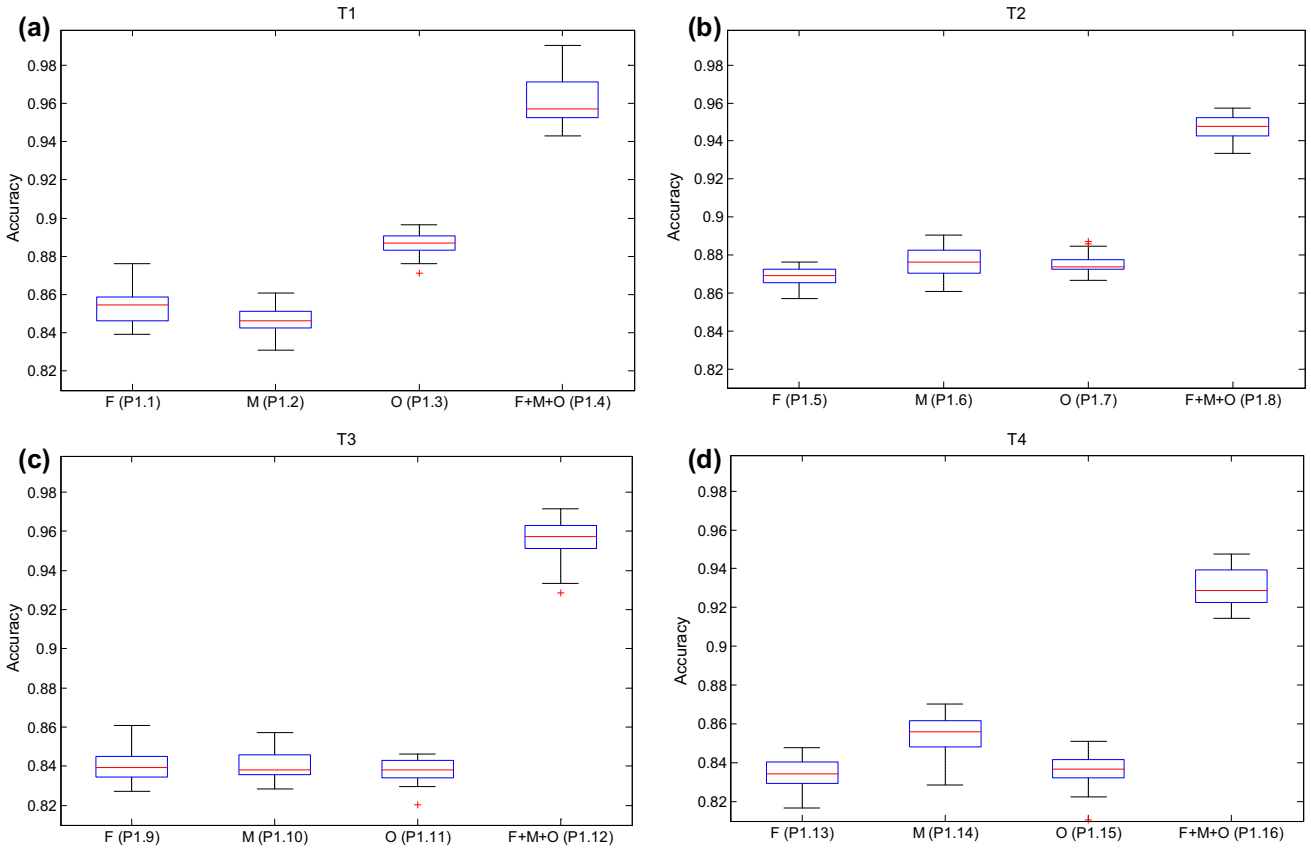

Fig. 6 Identification rates for matching tasks and different electrode regions: **a** Task1, **b** Task2, **c** Task3, **d** Task4

Table 4 Mean accuracies according to Protocol P2

Test set	Training set			
	T1R1 + T1R3 (%)	T2R1 + T2R3 (%)	T3R1 + T3R3 (%)	T4R1 + T4R3 (%)
T1R2	96.15	91.22	89.73	89.49
T2R2	96.44	94.72	86.48	91.45
T3R2	92.78	90.13	95.50	88.42
T4R2	95.01	94.91	87.12	93.10
TbEO	1.92	1.43	2.32	2.08
TbEC	3.45	1.45	1.92	2.44

not essential for achieving a good performance provided a nonresting task have been used for training and testing. In this investigation, data from multiple task types are pre-combined for classifier training. As shown in Fig. 7, the size of the training set is gradually increased by adding the training data coming from different task types (data from all nine electrodes from the three scalp regions were used). The identification accuracy increased steadily with the accumulation of more training data until the performance curve becomes flat. Under this experimental protocol, accuracy rates greater than 99 % have been achieved.

The results in P3 indicate that by concatenating different types of motor movement/imagery data for system training, the identification performance noticeably improved. However, this improvement appears to saturate as the training data volume is further increased.

4.2 Verification scenario

The results presented in this section are from the evaluation of the proposed system in the verification scenario. The detection error trade-off (DET) curves, which reveal the relationships between false acceptance rates (FAR) and false rejection rates (FRR) at different operating thresholds, have been used throughout this work to evaluate the proposed system [36]. In some cases, equal error rate (EER)

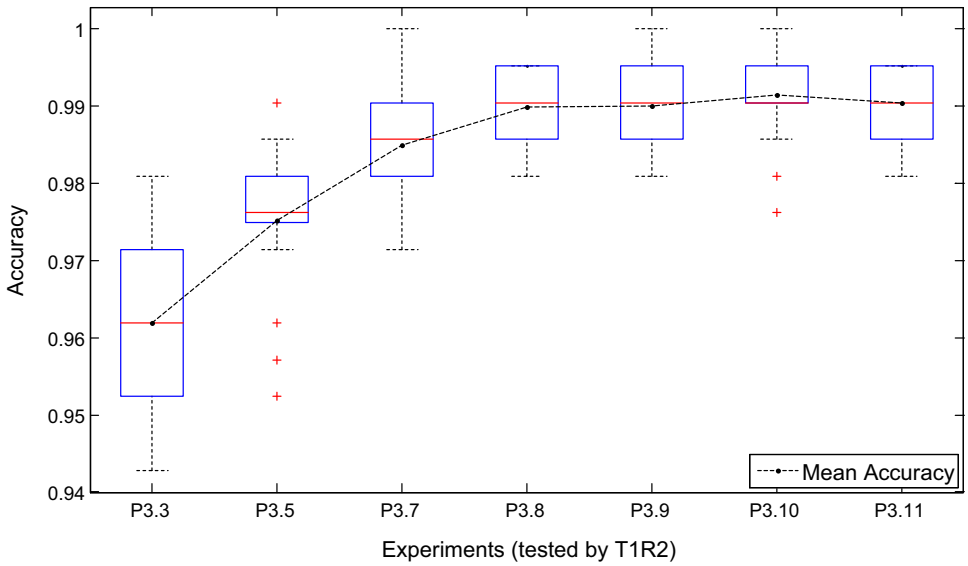

Fig. 7 Protocol P3—experimental results (The labels under the boxplots indicate the experiments as in Table 3)

has also been used for comparative analysis. The publicly available software employed here to generate DET curve was provided by the National Institute of Standards and Technology (NIST) [37]. Note that the Fisher's LDA was used for verification, through a series of binary classifications where each subject was in turn classified against the rest of the subjects in the database.

4.2.1 *Protocol P1: analysing the impacts of different electrode positioning*

The DET curves in Fig. 8 depict the verification performance of signals captured from three scalp regions. Data from matching tasks were used for training and testing. Only the results from task T1 are reported as an example to analyse the impact of electrode locations; results obtained from other tasks also exhibited similar trends. Like previous experiments, data of R1 and R3 together were used for training and the data from R2 were used for testing.

It can be observed from Fig. 8 that the occipital electrodes produced the best overall performance amongst the three scalp regions while the motor cortex electrodes produced the worst. The data obtained from occipital lobe and frontal region provided comparable EERs. Despite the results indicating that the occipital lobe seems to be a slightly better electrode location, the difference between the performances amongst these regions is quite small. Hence, the impact from the electrode locations is not conclusive.

4.2.2 *Protocol P2: the impact of nonmatching tasks*

The DET curves in Fig. 9 depict the performance when different types of tasks were used for testing the system

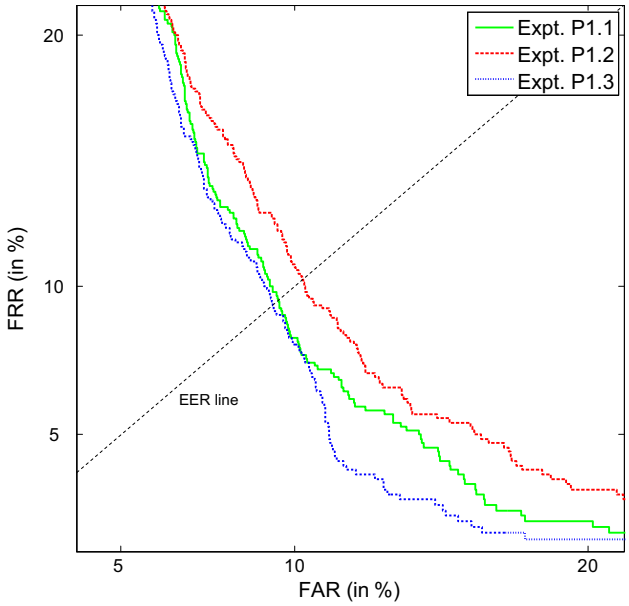

Fig. 8 DET curves for Task 1 for Experiments P1.1 (frontal lobe), P1.2 (motor cortex) and P1.3 (occipital lobe)

while it was trained by T1 as described in Protocol P2. Figure 9a shows the results obtained using data from the occipital lobe (three electrodes), since these electrodes produced the best overall performance in the previous experiments. For comparison purposes, the data from the two baseline tasks (TbEO and TbEC) were also used for testing. For matched training and test data, the EER was 8.26 %. For nonmatching motor movement/imagery tasks, the performances were found very similar. For example, the EERs were 8.09 % for T2R2O and 7.83 % for T4R2O, respectively, which were even lower than when it was tested by the same type task data (T1R2O). However, much

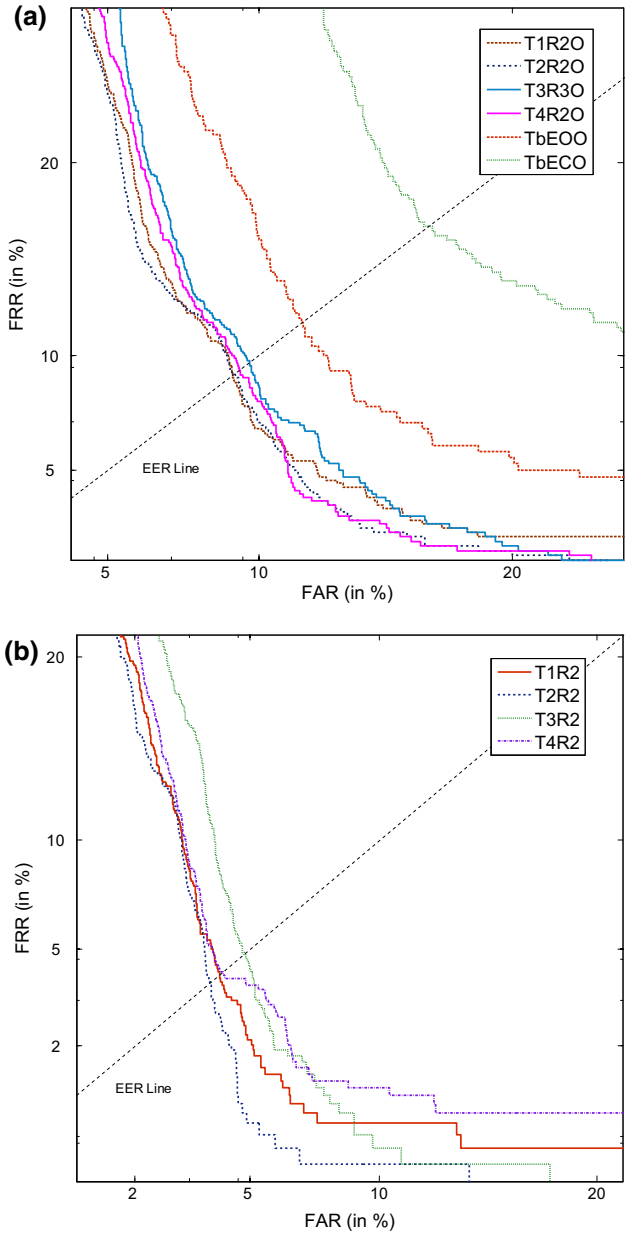

Fig. 9 DET curves showing the impact of testing with different task types when the system is trained with T1 **a** only occipital lobe data are used, **b** all nine electrodes are used (The legends indicate the test set)

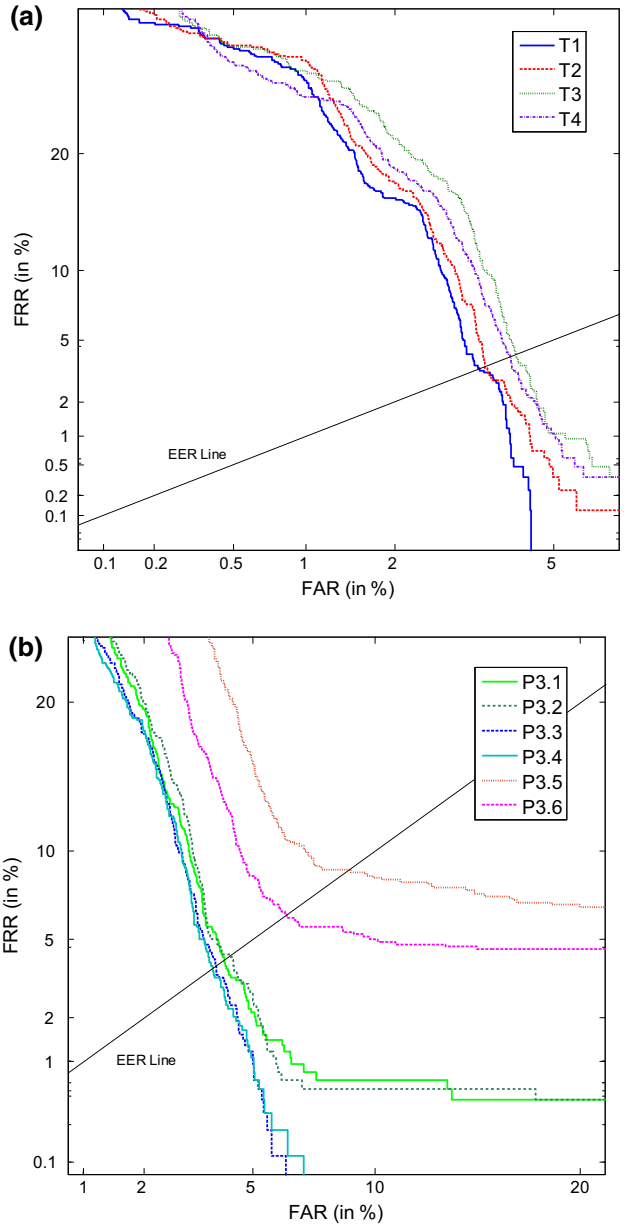

Fig. 10 **a** Tested with T1R2, cross-task increasing the training data size, **b** tested with one run of one task, trained with the rest of data for the four tasks

worse performances were exhibited by the two curves which represent the two baseline tasks. Furthermore, the TbEO curve indicates the performance when the system was tested using the eyes open baseline data, resulting in better verification rates than that provided by the data obtained while eyes were closed (TbEC).

In Fig. 9b, data from all the nine electrodes (in the three regions) was employed. The lowest system EER achieved was 2.785 % when it was tested with the data from task T2. The error rates in these sets of experiments have reduced by about three times when the number of electrodes employed was raised from three to nine. In this implementation, all the DET curves are again very compactly clustered.

It is therefore evident that when the system is trained by data from one motor movement/imagery task and tested by another, the verification performance does not necessarily deteriorate than that from the task-matching test. However, when tested by the data generated from the resting states, the performances degraded quite significantly.

4.2.3 Protocol P3: concatenating different task data for training

Extensive tests, using nine electrodes shown in Fig. 4, were conducted to investigate the effect of training using an aggregation of all the four tasks on verification rates. Of the available twelve same-length data recordings, only one recording has been set aside for testing and all the remainder were used for training. Figure 10a shows four DET curves, while four different tasks were used for testing: for example, the curve marked T1 indicates using data of one run from T1 for testing while the data of all the remaining runs (11 runs) for training. The results suggest that all the EERs fall within around 3–5 % by using the accumulated data.

The DET curves in Fig. 10b depict the results obtained from a subset of experiments in Protocol P3. Here, the features from different tasks were gradually concatenated to train the system. Data from T1R2 alone were used as the test set. The system achieved the lowest EER of 2.63 % by combining the data from T1 and T2 for training. Although the EER dropped initially with the addition of extra training data, after a point there was no further improvement in performance.

It is evident that the DET curves cluster in three groups depending on the volume of the training data. For the group with the smallest training set, using the data of single runs from a single task for training, the lowest performances were observed and the training data of R3 (P3.2) provide better results than R1 (P3.1). This performance variation also indicates that stable operation of EEG biometrics verification systems may not be possible with short amounts of training data. The performance improved with increasing the training data volume (from multiple data recordings/tasks) and alleviated the performance variation.

4.3 Comparative analysis

Table 5 shows some of the most recent works that relate to the proposed experiments. All of these reports of used one or both of two popular publicly available databases for EEG biometrics: the MM/I data set and the UCI VEP database [38]. The UCI VEP database contains comparable number of subjects with MM/I data set, but with only single recording session. Indeed, the results reported in [9] and [42] provided comparable identification performance when other factors (number of electrodes and subjects, for instance) are taken into consideration.

The proposed system provided comparable performance with the state-of-art systems in both identification and verification scenarios, but employed much less number of electrodes. Compared with most of the reports in Table 5, the proposed system separated the training and test data by different recordings, which is a step further towards the realistic biometric scenarios.

5 Conclusions

In this paper, we have explored the impact of user activity on the performance of an EEG-based biometric system using wavelet features. Using EEG biometric signals based on the time derivative of wavelet coefficients, we investigated impact of electrode placements and the type and quantity of training data on the system accuracy using a mixture of motor movement/mental imagery tasks. We constructed three protocols to verify the questions raised in Sect. 1, mainly aimed at establishing the impact of training strategies and data volume on performance.

The results indicate that for the proposed experimental design there is no clear difference in performance amongst scalp regions. It was also found that the recognition performance was not sensitive to nonmatching motor movement/imagery tasks used for training and testing. Aggregated EEG data obtained from different types of user activity from separate recordings was explored to build more robust training models. Results clearly indicate that increasing the training data volume, irrespective of the type

Table 5 Comparison with related works

Reports	Features	Task(s)	Electrode(s)	Run(s)	Subjects	Performance (%)
Motor movement/imagery data set (MM/I data set)						
Rocca et al. [9]	PSD & spectral coherence	EO & EC	56	1	108	CRR: 100
Fraschini et al. [39]	Eigenvector centrality	EO & EC	64	1	109	EER: 4.4
Proposed work	Wavelet coefficients	T1–T4	9	3	108	CRR: 99
						EER: 4.5
UCI EEG database data set (VEP data set)						
Su and Farzin [40]	EEMD-based InsAmp	Visual	1	1	118	CRR: 95.9
Yazdani et al. [41]	AR coefficients & PSD	Visual	64	1	20	CRR: 100
Brigham et al. [42]	AR coefficients	Visual	64	1	120	CRR: 98.96
Huang et al. [43]	Root-mean-square values	Visual	64	1	116	CRR: 95.1

of activity used, improves identification and verification performance.

The overall conclusion is that there appears to be substantial flexibility in the choice of user activity employed for training and testing such systems. The work has also indicated that data from different types of motor movement/imagery activity may be aggregated to provide more robust training of the system without any adverse effects. This flexibility with regards to types of user activity could result in systems that are easier to develop, deploy and use in a range of applications. Future work will be focused on evaluating the robustness of this approach when collecting data with long time intervals between training and testing as well as data from low-cost EEG sensors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Jain AK, Pankanti S, Prabhakar S, Hong L, Ross A (2004) Biometrics: a grand challenge. In: Proceedings twentieth international conference pattern recognition, vol 2, pp 935–942
2. Kawabata N (1973) A nonstationary analysis of the electroencephalogram. IEEE Trans Biomed Eng 20(6):444–452
3. Poulos M, Rangoussi M, Chrissikopoulos V, Evangelou A (1999) Person identification based on parametric processing of the EEG. In: Proceeding of sixth international conference electronics, circuits, and systems vol 1, pp 283–286
4. Poulos M, Rangoussi M, Chrissikopoulos V, Evangelou A (1999) Parametric person identification from the EEG using computational geometry. In: Proceedings of sixth international confernce electronics, circuits, and systems vol 2, pp 1005–1008
5. Poulos M, Rangoussi M, Alexandris N (1999) Neural network based person identification using EEG features. In: International conference acoustics, speech, and signal processing, vol 2, pp 1117–1120
6. Polich J (2007) Updating P300: an integrative theory of P3a and P3b. Clin Neurophysiol 118(10):2128–2148
7. Su F, Xia L, Cai A, Wu Y, Ma J (2004) EEG-based personal identification: from proof-of-concept to a practical system. In: Proceedings twentieth international conference pattern recognition, pp 3728–3731
8. Lee HJ, Kim HS, Park KS (2013) A study on the reproducibility of biometric authentication based on electroencephalogram (EEG). Proceedings of sixth international conference neural engineering (NER), pp 13–16
9. La Rocca D, Campisi P, Vegso B, Cserti P, Kozmann G, Babiloni F, Fallani FDV (2014) Human brain distinctiveness based on EEG spectral coherence connectivity. IEEE Trans Biomed Eng 9294:1–7
10. Luck SJ (2014) An introduction to the event-related potential technique. MIT Press, Cambridge
11. Donchin E, Spencer KM, Wijesinghe R (2000) The mental prosthesis: assessing the speed of a P300-based brain-computer interface. IEEE Trans Rehabil Eng 8(2):174–179
12. Singhal GK, RamKumar P (2007) Person identification using evoked potentials and peak matching. In: Proceedings of conference biometrics symposium, pp 1–6
13. Yearn SK, Suk HI, Lee SW (2013) EEG-based person authentication using face stimuli. In: International winter workshop brain-computer interface (BCI), pp 58–61
14. Palaniappan R, Raveendran P (2002) Individual identification technique using visual evoked potential signals. Electron Lett 38(25):1634–1635
15. Snodgrass JG, Vanderwart M (1980) A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol 6(2):174
16. Palaniappan R, Ravi KVR (2003) A new method to identify individuals using signals from the brain. In: Proceedings of joint conference of information, communications and signal processing, and fourth pacific rim conference on multimedia, vol 3, pp 1442–1445
17. Palaniappan R (2007) EEG based biometric framework for automatic identity verification. J VLSI Sig Proc 49:243–250
18. Palaniappan R, Mandic DP (2007) Biometrics from brain electrical activity: a machine learning approach. IEEE Trans Pattern Anal Mach Intell 29(4):738–742
19. Palaniappan R (2005) Identifying individuality using mental task based brain computer interface. In: International conference intelligent sensing and information processing, pp 238–242
20. Marcel S, Millan JDR (2007) Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. IEEE Trans Pattern Anal Mach Intell 29(4):743–752
21. Chuang J, Nguyen H, Wang C, Johnson B (2013) I think, therefore I am: usability and security of authentication using brainwaves. In: Financial cryptography and data security. Springer, Berlin, pp 1–16
22. NeuroSky, main website http://neurosky.com/. Accessed 2 May 2015
23. Kostilek M, St'astny J (2012) EEG biometric identification: repeatability and influence of movement-related EEG. In: Proceedings of international applied electronics (AE), pp 147–150
24. Ghitza O, Sondhi MM (1993) Hidden Markov models with templates as non-stationary states: an application to speech recognition. Comput Speech Lang 7(2):101–119
25. Tarvainen MP, Hiltunen JK, Ranta-aho PO, Karjalainen PA (2004) Estimation of nonstationary EEG with Kalman smoother approach: an application to event-related synchronization (ERS). IEEE Trans Biomed Eng 51(3):516–524
26. Mallat S (1999) A wavelet tour of signal processing. Academic Press, CA, USA
27. Fugal DL (2009) Conceptual wavelets in digital signal processing: an in-depth, practical approach for the non-mathematician. Space and Signals Technical Publishing, CA, USA
28. Daubechies Ingrid (1992) Ten lectures on wavelets, vol 61. Society for Industrial and Applied Mathematics, Philadelphia
29. Yang S, Deravi F (2013) Wavelet-based EEG pre-processing for biometric applications. In: Fourth international conference emerging security technologies (EST), pp 43–46
30. Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR (2004) BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans Biomed Eng 51(6):1034–1043
31. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE (2000) Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. Circulation 101(23):e215–e220

32. EEG Motor Movement/Imagery Dataset, http://www.physionet.org/pn4/eegmmidb/. Accessed 1 May 2014
33. Homan RW, Herman J, Purdy P (1987) Cerebral location of international 10–20 system electrode placement. Electroencephalogr Clin Neurophysiol 66(4):376–382
34. Tortora GJ, Derrickson BH (2011) Principles of anatomy and physiology. Wiley, Chichester
35. He SQ, Dum RP, Strick PL (1995) Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. J Neurosci 15(5):3284–3306
36. Martin A, Doddington G, Kamm T, Ordowski M, Przybocki M (1997) The DET curve in assessment of detection task performance. National inst of standards and technology Gaithersburg MD
37. DET-Curve Plotting software for use with MATLAB, http://www.itl.nist.gov/iad/mig/tools/DETware_v2.1.targz.htm. Accessed 2 May 2015
38. UCI Machine Learning Repository: EEG database data set. [Online]. https://archive.ics.uci.edu/ml/datasets/EEG+Database. Accessed 22 Apr 2015
39. Fraschini M, Hillebrand A, Demuru M, Didaci L, Marcialis G (2015) An EEG-based biometric system using eigenvector centrality in resting state brain networks. IEEE Signal Process Lett 22(6):666–670
40. Yang S, Deravi F (2014) Novel HHT-based features for biometric identification using EEG signals. In: IEEE of 22nd international conference on pattern recognition (ICPR), pp 1922–1927
41. Yazdani A, Roodaki A, Rezatofighi SH, Misaghian K, Setarehdan SK (2008) Fisher linear discriminant based person identification using visual evoked potentials. In: 9th international conference on signal processing (ICSP), pp 1677–1680
42. Brigham K, Vijaya Kumar BVK (2010) Subject identification from electroencephalogram (EEG) signals during imagined speech." In: Fourth IEEE international conference on biometrics: theory applications and systems (BTAS), pp 1–8
43. Huang X, Altahat S, Tran D, Sharma D (2012) Human identification with electroencephalogram (EEG) signal processing. In: International symposium on communications and information technologies (ISCIT), pp 1021–1026
44. Campisi P, La Rocca D (2014) Brain waves for automatic biometric based user recognition. In: IEEE transactions on information forensics and security, pp 782–800