Pattern Recognition Letters
journal homepage: www.elsevier.com

Learning Deep Features for Task-Independent EEG-based Biometric Verification

Emanuele Maiorana[a,**]

[a]Department of Engineering, Roma Tre University, Via V. Volterra 62, 00146, Rome, Italy

ABSTRACT

Considerable interest has been recently devoted to the exploitation of brain activity as biometric identifier in automatic recognition systems, with a major focus on data acquired through electroencephalography (EEG). Several researches have in fact confirmed the presence of discriminative characteristics within brain signals recorded while performing specific mental tasks. Yet, to make EEG-based recognition appealing for practical applications, it would be highly advisable to investigate the existence and permanence of such distinctive traits while performing several different mental tasks. In this regard, the present study evaluates the feasibility of performing task-independent EEG-based biometric recognition. A deep learning approach using siamese convolutional neural networks is employed to extract, from the considered EEG recordings, subject-specific template representations. An extensive set of experimental tests, performed on a multi-session database comprising EEG data acquired from 45 subjects while performing six different tasks, is employed to evaluate whether it is actually possible to verify the identity of a subject using brain signals, regardless the performed mental task.

1. Introduction

Biometric traits employed for automatic recognition purposes have been traditionally categorized into physical and behavioral data. Alongside these two types, cognitive biometric traits have been recently proposed to design biometric recognition systems (Revett, 2012). This latter kind of identifiers involves the acquisition of biosignals generated by the nervous system in response to a stimulus or during a task. Notable examples belonging to this group include the heart activity (da Silva Luz et al., 2018), the skin electrodermal activity (Bianco and Napoletano, 2019), and also the brain activity (Gui et al., 2019). Resorting to these traits offers several advantages with respect to the use of data such as fingerprint, iris, and face. In fact, the aforementioned activities cannot be recorded at a distance, thus making it hard to covertly capture them and perform presentation attacks. Using cognitive biometric data for recognition purposes also inherently solves liveness detection issues. Furthermore, systems relying on such traits may perform continuous recognition, and provide robustness against coercive attacks, since stress conditions can be easily detected analyzing the collected data without requiring additional hardware.

While recognition systems based on heart activity have been extensively investigated, with recent applications proposed even for wearable devices (Pinto et al., 2018), research on the usage of brain activity for biometric recognition purposes is still on the rise, with several aspects to be properly investigated and developed before brain-based biometric systems could be deployed in practical applications. In more detail, solutions available to monitor brain activity include functional magnetic resonance imaging (fMRI), near infrared spectroscopy (NIRS), positron emission tomography (PET), magnetoencephalography (MEG), and electroencephalography (EEG). Among them, only this latter can be considered in the context of biometric recognition, due to the relative inexpensiveness of the associated acquisition devices and their ease of use.

EEG signals are generated by sensing, on a subject's scalp, the electric field whose characteristics depend on the firing of spatially-aligned cortex pyramidal neurons. Such brain activity is commonly categorized into five main oscillatory rhythms, namely δ (0.5 ÷ 4Hz), θ (4 ÷ 8Hz), α (8 ÷ 13Hz), β (13 ÷ 30Hz), and γ (> 30Hz) (Niedermeyer and Da Silva, 2005). Current EEG sensing technology is among the major issues limiting the use of EEG signal as biometric identifiers. In fact, either wet electrodes are used, with the consequent requirement for elec-

**Please cite this work as: Emanuele Maiorana, "Learning Deep Features for Task-Independent EEG-based Biometric Verification," Pattern Recognition Letters, Vol. 143, pp: 122-129,March 2021. Digital Object Identifier 10.1016/j.patrec.2021.01.004.

trolyte gel to improve conductivity, or dry electrodes can be employed, with their shape however causing ache for prolonged acquisitions. In both cases, there is therefore the need to limit the amount of channels in the recording devices, in order not to make them too uncomfortable for the users (Pinegger et al., 2016). Moreover, novel approaches are required to extract discriminative features from the collected EEG signals, with the aim of performing recognition achieving error rates comparable with those granted by well-established biometric traits. In this regard, it has to be remarked that the vast majority of investigations conducted in this area have focused on the search for distinctive features in brain signals acquired according to specific protocols, such as those involving resting states, imagined movements, speech imagery, and so forth (Yang and Deravi, 2017). However, in order to support the exploitation of brain signals as biometric identifier in real-life practical applications, it would be instead highly relevant to investigate the existence of stable subject-specific characteristics in EEG signals collected while performing different mental tasks. This is actually the scope of the present paper, which evaluates the feasibility of performing task-independent EEG-based biometric recognition resorting to deep learning approaches. In more detail, the contributions of the present study are the following ones:

- the existence of EEG task-independent discriminative features, allowing to perform EEG-based biometric recognition regardless the performed mental task, is here evaluated. In order to perform such evaluation, a multi-session database, comprising EEG data acquired from 45 subjects while performing six different tasks during five sessions spanning an overall period of more than one year, is here exploited;
- in order to explicitly learn EEG characteristics stable across the considered mental tasks, an appropriate training strategy relying on siamese convolutional neural networks (CNNs) is here adopted;
- the use of different models for distinct EEG channels is here proposed, with the purpose of learning specific discriminative features for each area of the brain;
- the recognition capabilities of EEG signals recorded while performing different mental tasks are here evaluated comparing recordings taken at either short and long time distances between enrolment and verification, to evaluate the longitudinal behavior of the estimated discriminative characteristics;
- an analysis of the recognition performance achievable when reducing the number of employed EEG channels, in order to improve the usability of the proposed system, is also performed.

The current state of the art on automatic recognition systems using brain signals as biometric identifiers is provided in Section 2. The approach here proposed to derive discriminative representations from the considered EEG recordings, exploiting CNNs trained with a siamese strategy, is outlined in Section 3. The recognition results achieved in tests performed in verification modality, exploiting the availability of an EEG database comprising samples taken from 45 subjects during five recording sessions, each consisting of six different tasks, are then discussed in Section 4, while conclusions are finally drawn in Section 5.

2. State of the Art on EEG-based Biometric Recognition

Although biometric recognition based on brain activity has been postulated in the early '80s (Stassen, 1980), a systematic investigation on the use of EEG signals as biometric identifiers has been conducted only in the last decade (Gui et al., 2019). In order to extract discriminative information, EEG recordings are typically first preprocessed with temporal and spatial filters to reduce the amount of artifacts not related to brain activity, like those associated to endogenous factors such as eye blinking and muscular activity, or to exogenous sources such as power supply noise (Yang and Deravi, 2017).

Traditionally, hand-crafted features are then extracted from the obtained signals. Most of the proposed approaches rely on representations derived by treating separately signals acquired with different electrodes, resorting for instance to autoregressive (AR), wavelet-based, power spectral density (PSD), or mel-frequency cepstrum coefficients (MFCCs) modeling for each channel (Campisi and La Rocca, 2014). Even methods estimating functional brain connectivity, therefore focusing on temporal dependencies among EEG signals generated in different brain areas, have been proposed using measures such as correlation, spectral coherence, and Granger causality to generate biometric templates from EEG data (Friston, 2011).

More recently, also deep learning approaches have been exploited to define discriminative representations of EEG signals. Convolutional neural networks (CNNs) have been applied to brain signals to perform recognition for the first time using a shallow network with two layers (Ma et al., 2015). Frameworks comprising adversarial CNNs and recurrent neural networks have been proposed too (Ozdenizci et al., 2019; Maiorana, 2020).

Regarding the performance achieved in literature, several studies have claimed to obtain perfect recognition accuracy when using EEG signals as biometric identifiers (Chen et al., 2016). Unfortunately, the reliability of many of such works is often undermined by a common misconduct, that is, the execution of experimental tests on databases comprising EEG data collected during a single recording session for each considered subject (Ruiz-Blondet et al., 2016). Under this scenario, the estimated performance may depend more on session-specific recording conditions than on individual characteristics of the involved subjects (Ozdenizci et al., 2019). For this reason, proper testing of an EEG-based biometric recognition system should be instead performed using multi-session datasets, comparing signals recorded in different days. When such conditions are considered, the achievable recognition rates notably worsen with respect to tests in which data from the same session are employed for both enrolment and recognition purposes, motivating the need for longitudinal studies on EEG discriminative capabilities (Maiorana and Campisi, 2018).

Single-session datasets have been considered in almost all the works evaluating EEG-based task-independent recognition, as in (Fraschini et al., 2019) where data from the PhysioNet database (Goldberger et al., 2000), comprising samples collected from users performing 14 tasks during a single session, have been exploited. As in (Fraschini et al., 2019), EEG representations based on functional connectivity have been exploited in (Wang et al., 2019), where tests have been conducted on four tasks of the PhysioNet database, and on an in-house database comprising single-session recordings too. Nonetheless, the results in (Fraschini et al., 2019) and (Wang et al., 2019) show a degradation in recognition performance in cross-task scenarios, with respect to the within-task test conditions.

Databases collected during multiple sessions have been instead considered in (Kong et al., 2018) and (Kumar et al., 2019). EEG signals from nine users performing four tasks during two sessions have been used in the former paper, while data captured from 30 subjects performing four tasks during three sessions have been exploited in the latter. However, tests in both (Kong et al., 2018) and (Kumar et al., 2019) have been performed using acquisitions from all the available sessions for enrolment purposes, thus still adopting improper experimental conditions.

Distinct sessions have been instead properly considered for enrolment and recognition purposes in (Del Pozo-Banos et al., 2018) and (Vinothkumar et al., 2018). In both these studies, notable performance losses have been actually observed when comparing EEG signals recorded in different sessions, in addition to the worsening due to the cross-task scenario. It has yet to be remarked that EEG acquisitions from only five subjects, performing five tasks in two different days, have been considered in (Del Pozo-Banos et al., 2018), while only 15 subjects carrying out five tasks during two recording sessions have been taken into account in (Vinothkumar et al., 2018). The number of users available in (Del Pozo-Banos et al., 2018) and (Vinothkumar et al., 2018) is therefore too low, even for studies regarding EEG-based biometric recognition, to derive meaningful conclusions regarding the observed behaviors. Conversely, as it will be detailed in Section 4, the present study has been conducted on a much larger database, comprising EEG recordings taken from 45 subjects during a period lasting more than one year, therefore representing a much more reliable basis to derive meaningful conclusions regarding the existence of EEG cross-task discriminative characteristics.

It has also to be mentioned that the approaches in (Del Pozo-Banos et al., 2018) and (Vinothkumar et al., 2018), like all the others so far employed to investigate the existence of task-independent characteristics in brain signals, have resorted to hand-crafted features to generate the EEG representations employed in the proposed classifiers. A supervised deep-learning-based approach is instead adopted in the present study, with the aim of explicitly learning EEG characteristics stable across the considered mental tasks, through an appropriate training strategy. Specifically, as outlined in Section 3, siamese CNNs are trained to accomplish such aim, performing biometric recognition on a multi-session and multi-task EEG database.

Furthermore, scenarios involving both short- and long-term distances between enrolment and verification are here taken into account, thus providing a longitudinal analysis of the employed EEG discriminative characteristics. Tests aimed at verifying the effectiveness of the proposed solution when using a limited number of EEG channels, in order to improve the usability of the proposed system, are also here discussed.

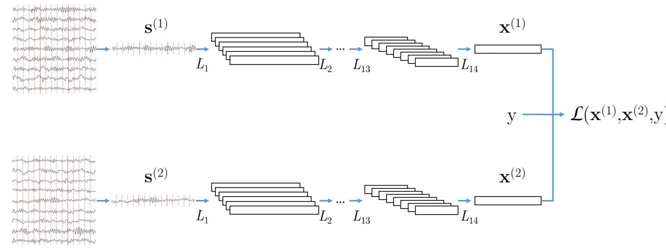


Fig. 1. Proposed channel-specific siamese network training.

3. Proposed Biometric Verification System

The approach here employed to derive discriminative features from EEG traits stems from the method proposed in (Maiorana, 2019), where deep learning has been applied for the first time to brain signals for the implementation of a biometric verification system. Specifically, siamese CNNs have been there trained over multi-session EEG data recorded in eyes-closed (EC) resting states.

As shown in Figure 1, a siamese network uses two or more identical subnetworks, with the same architecture and sharing the same parameters and weights, simultaneously updated at each step of the learning process. The loss evaluated for back-propagation purposes depends on the Euclidean distance of the representations generated by the employed subnetworks. In more detail, the computed loss should be minimized trying to lower the distance between representations derived from input samples belonging to the same class, while increasing the distance between representations obtained from inputs belonging to different classes. Indicating as $\mathbf{s}^{(1)}$ and $\mathbf{s}^{(2)}$ the samples fed to the two subnetworks composing the considered siamese framework, and as $D_{\mathbf{x}^{(1)},\mathbf{x}^{(2)}}$ the Euclidean distance computed between the generated representations $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, a contrastive loss function is evaluated during the performed training as

$$\mathcal{L}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, y) = (1-y)\frac{1}{2}D^2_{\mathbf{x}^{(1)},\mathbf{x}^{(2)}} + y\frac{1}{2}[max(0, d-D_{\mathbf{x}^{(1)},\mathbf{x}^{(2)}}]^2, \quad (1)$$

being y the label associated with the considered pair of inputs, with $y = 0$ for samples from the same class and $y = 1$ otherwise. The margin d is employed to control which pairs with samples from distinct classes should contribute to the loss and the learning process.

The tests performed in (Maiorana, 2019) have highlighted that, to define EEG templates to be used in verification modality, it is preferable to exploit siamese networks with raw time-dependent EEG signals as inputs, rather than first extracting hand-crafted features as suggested for identification purposes in (Maiorana, 2020). The same approach is therefore here employed, designing an end-to-end learning process with CNNs directly applied to the available EEG signals.

With respect to (Maiorana, 2019), an attempt at providing greater generalization is here proposed and investigated, training distinct CNN models for each individual EEG channel, instead of learning a single network for all the signals collected through the employed electrodes as done in (Maiorana, 2019). This choice aims at learning discriminative characteristics peculiar of each area of the brain, trying in this way to improve the achievable recognition rates.

The processing pipeline characterizing the enrolment and verification stages of the proposed EEG-based recognition system is detailed in the following Section 3.1.

3.1. Employed Processing Pipeline

It is assumed that EEG data are collected using C electrodes placed on the subjects' scalp. Given the outcomes of literature researches such as (Maiorana et al., 2016a) and (Maiorana and Campisi, 2018), EEG signals are filtered during preprocessing to retain the frequencies which have shown to contain the most discriminative and permanent EEG content, that is, the subband $[\alpha, \beta] = 8 \div 30$Hz. Without discarding significant information, thanks to the previous filtering operation and given the Nyquist–Shannon sampling theorem, a downsampling to $S = 64$Hz is then performed to reduce the computational complexity of the subsequent processing. Lowering the sampling frequency in fact allows to define shorter sequences as inputs to the employed CNNs, as detailed in the following. A spatial common average referencing (CAR) filter is also employed to limit the effects of possible incorrect reference positioning. No further preprocessing is applied to the EEG signals employed in the tests reported in the following. Actually, techniques defined to remove artifacts from EEG recordings (Maiorana et al., 2016b) have been considered during the performed tests, yet no significant difference in recognition performance has been obtained when applying them to the employed signals, and they have been consequently omitted from the proposed system in order to reduce the required computational cost.

Following the approaches already employed in several studies on EEG-based biometric recognition (Maiorana et al., 2016a; Maiorana and Campisi, 2018), the obtained signals are then segmented into frames lasting $H = 5$s, a duration which should guarantee a proper trade-off between the requirement for having a manageable sample size, and the need for having in each frame enough information for reliably estimating stationary features. An 80% overlap factor between consecutive frames is also employed to generate a number of samples allowing to properly train the employed CNNs. A single frame is considered as either an enrolment or verification sample in the proposed system, and comprises C temporal sequences $\mathbf{s}_c$, $c = 1, \ldots, C$, each having size $1 \times S \cdot H = 1 \times 320$. The obtained unidimensional signals are fed as input to C distinct CNNs, sharing the same framework reported in Table 1, yet trained over data recorded by different electrodes to capture channel-specific discriminative characteristics. The employed network is the one giving the best results among a large set of tested configurations, including those in (Maiorana, 2020), where different choices have been made regarding the size of the employed convolutional filters and the adoption of max-pooling and dropout (DO) layers. Nonetheless, no claim of optimality is here made. A set of C representations $\mathbf{x}_c$, $c = 1, \ldots, C$, is therefore derived for each EEG frame.

The verification process is carried out by separately comparing corresponding channels of frames collected during enrolment and authentication phases, thus generating C different scores for each verification probe. Score fusion is then applied to generate a single outcome for the verification process of a frame. Thresholding is finally performed to take a decision on the identity of the subject whose EEG frame has been analyzed.

Table 1. Employed CNN. Conv layers comprise a batch normalization step.

#	Layer	Filter	Pad	Input	Output
L_1	Conv	(1×5×1)×16	[0,2]	1×320×1	1×320×16
L_2	ReLu	-	-	1×320×16	1×320×16
L_3	MP	1×3	-	1×320×16	1×106×16
L_4	Conv	(1×5×16)×32	[0,2]	1×106×16	1×106×32
L_5	ReLu	-	-	1×106×32	1×106×32
L_6	MP	1×3	-	1×106×32	1×35×32
L_7	Conv	(1×3×32)×64	-	1×35×32	1×33×64
L_8	Relu	-	-	1×33×64	1×33×64
L_9	MP	1×3	-	1×33×64	1×11×64
L_{10}	Conv	(1×3×64)×128	-	1×11×64	1×9×128
L_{11}	ReLu	-	-	1×9×128	1×9×128
L_{12}	MP	1×3	-	1×9×128	1×3×128
L_{13}	DO	-	-	1×1×128	1×3×128
L_{14}	Conv	(1×3×128)×256	-	1×3×128	1×1×256

Comparison scores are generated employing, for each subject, models trained with one-class support vector machines (SVMs) over representations of each channel derived from enrolment frames, instead of computing Euclidean distances between enrolment and verification samples as in (Maiorana, 2019). It is worth remarking that the proposed processing pipeline has been designed in order to make the entire procedure, including the feature extraction process and the structure of the employed CNNs, independent on the number of employed channels C, thus leaving the possibility of keeping the recognition process unchanged as the number of employed electrodes varies.

4. Experimental Tests

In order to evaluate whether task-independent EEG discriminative characteristics exist, and whether the proposed approach based on siamese CNNs could learn them, an extensive set of experimental tests has been carried out on the largest multi-session dataset, in terms of considered subjects and employed elicitation protocols, ever exploited to perform EEG-based biometric recognition. Specifically, the database here considered has been also used in (Maiorana et al., 2016a), (Maiorana and Campisi, 2018), and (Maiorana, 2020), and comprises EEG data collected at an original sampling rate of 256 Hz using a GALILEO BE Light amplifier with 19 wet electrodes placed according to the 10-20 International system, shown in Figure 2. Five different recording sessions, indicated as R_i, $i = 1, \ldots, 5$, have been performed for each of 45 healthy subjects, taking EEG acquisitions with the involved subjects comfortably seated on a chair in a quiet and dimly lit room, in order to reduce artifacts due to stress or distractions. While the first three sessions have been taken in the time span of a month, an average of six and fifteen months have been then passed before the fourth and fifth sessions, respectively, covering an overall period of more than one year between the first and the last recording ses-

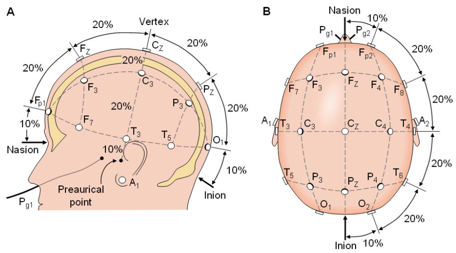


Fig. 2. The 10-20 International system seen from left (A) and above the head (B). The letters F, T, C, P, and O stand for frontal, temporal, central, parietal, and occipital lobes. (Jaakko Malmivuo and Robert Plonsey, Bioelectromagnetism, Oxford University Press, 1995, WEB version).

sions. Detailed information about the temporal distributions of the performed acquisitions are given in (Maiorana, 2020).

In (Maiorana et al., 2016a) and (Maiorana and Campisi, 2018) it has been shown that, although aging actually affects EEG characteristics, negligible consequences on recognition performance are observed when comparing signals captured within time lapses in the order of a month, while relevant effects are observable for larger periods. The availability of multiple recording sessions has been therefore exploited performing tests comparing EEG signals captured at either short time distance (*STD*) and long time distance (*LTD*). The former scenario is referred to comparisons among data captured at a time distance lower than one month, the latter one at distances of about 15 months. Furthermore, the usefulness of performing enrolment considering EEG data captured in more than one occasion has been also analyzed, computing recognition rates achievable with both single-session enrolment (*SSE*) and multiple-session enrolment (*MSE*). The experimental results referred to each of the four considered scenarios have been obtained performing comparisons between enrolment and verification data taken from the sessions summarized in Table 2. For instance, the recognition performance achievable when comparing signals taken at a long time distance, and using multiple sessions for enrolment, has been estimated averaging the results obtained when comparing data from R_1 and R_2 with recordings in R_5, data in R_1 and R_3 with those in R_5, and signals from R_2 and R_3 with acquisitions from R_5.

Each subject has performed six different mental tasks during each recording session, stimulating different brain activation patterns (Niedermeyer and Da Silva, 2005):

- resting state with eyes-closed (EC), tipically corresponding to a predominant α contribution in the parieto-occipital region of the scalp;
- resting state with eyes-open (EO), characterized by α desynchronization and increase in θ power with respect to EC;
- motor imagery (MI), with subjects asked to perform imaginary movements of left and right arms and legs;
- speech imagery (SI), with subjects asked to mentally reproduce the sound of a vowel observed on the screen. Both SI and MI tasks involve increased oscillatory activity in β band by the motor cortex in the centro-lateral side of the scalp;
- visual stimulation (VS), during which eight different geometric shapes are randomly shown on a screen;

Table 2. Data and comparisons evaluated in each considered test condition.

Time Distance	Enrolment Type	Enrolment Session(s)	Verification Session
STD	*SSE*	R_1	R_2
		R_1	R_3
		R_2	R_3
	MSE	R_1, R_2	R_3
LTD	*SSE*	R_1	R_5
		R_2	R_5
		R_3	R_5
	MSE	R_1, R_2	R_5
		R_1, R_3	R_5
		R_2, R_3	R_5

- mathematical calculation (MC), with subjects performing mathematical operations such as sums and differences. Both VS and MC implies greater involvement of the frontal lobe, with an increase in δ, β, and γ bands.

The acquired EEG data are treated as continuous streams to which the processing described in Section 3.1 is applied.

Tests have been first performed to verify the superiority of the proposed deep learning approach for EEG-based biometric verification over the use of hand-crafted features, as well as the effectiveness of looking for area-specific discriminative characteristics by separately modeling different channels, as described in Section 4.1. The recognition rates achievable under cross-task recognition scenarios are then outlined in Section 4.2, while Section 4.3 reports the results of a usability analysis, where the number of employed channels is limited in order to assess the recognition performance attainable when trying to reduce subject inconvenience.

4.1. *Effectiveness of Channel-specific Modeling*

In order to properly test the proposed verification system based on siamese networks in open-set conditions, a cross-validation experimental design has been carried out evaluating each considered scenario by dividing, for five iterations, the available subjects into two disjoint subsets: a training dataset comprising 30 subjects, and a testing dataset with the remaining 15 subjects, upon which recognition rates are evaluated.

To represent state-of-the-art EEG recognition systems based on hand-crafted features, fusion of AR and MFCC representations have been exploited as in (Maiorana and Campisi, 2018). Table 3 reports the recognition performance obtained on testing datasets when resorting to hand-crafted representations, to deep representations obtained as in (Maiorana, 2019), and to the approach here proposed, relying on separate CNN modeling for each of the $C = 19$ available channels, for all the six considered mental tasks. The results in Table 3 have been therefore obtained training CNNs on EEG data recorded while performing a single task, and using the obtained models to generate representations from signals acquired in the same conditions.

In more detail, siamese training has been performed using pairs of EEG frames taken, for each subject and for each considered protocol, from recordings of different sessions. Specifically, the first three sessions of each user, together with the

Table 3. Performance (EER, in %) comparison, in task-dependent conditions with $C = 19$, between EEG-based systems using representations based on: (a): hand-crafted features; (b): features learned using the same CNN model for all the employed channels as in (Maiorana, 2019); (c): features learned using channel-specific CNN models as here proposed. A single frame lasting 5s is used as verification probe. Best results highlighted in bold.

Test Condition		Eyes-closed (EC)	Eyes-open (EO)	Motor imagery (MI)	Speech imagery (SI)	Visual stimulation (VS)	Mathemat. calculation (MC)
		Protocol					
(a): Hand-crafted representations (AR+MFCC)							
SSE	*STD*	15.2 ± 0.3	21.4 ± 0.4	20.3 ± 0.4	18.7 ± 0.3	20.7 ± 0.4	20.7 ± 0.5
SSE	*LTD*	17.5 ± 0.2	25.8 ± 0.5	23.7 ± 0.4	23.5 ± 0.4	20.9 ± 0.4	23.1 ± 0.5
MSE	*STD*	12.1 ± 0.2	16.8 ± 0.3	16.7 ± 0.3	16.0 ± 0.3	17.6 ± 0.4	17.1 ± 0.4
MSE	*LTD*	15.6 ± 0.2	22.9 ± 0.4	22.4 ± 0.5	21.7 ± 0.4	19.4 ± 0.3	21.0 ± 0.4
(b): Deep representations obtained as in (Maiorana, 2019)							
SSE	*STD*	13.5 ± 0.3	13.1 ± 0.3	12.8 ± 0.3	12.6 ± 0.3	14.2 ± 0.3	14.4 ± 0.3
SSE	*LTD*	14.4 ± 0.3	15.4 ± 0.3	15.9 ± 0.3	16.1 ± 0.4	14.3 ± 0.3	14.6 ± 0.3
MSE	*STD*	9.6 ± 0.2	10.3 ± 0.2	11.2 ± 0.3	11.2 ± 0.3	11.6 ± 0.3	11.3 ± 0.2
MSE	*LTD*	10.1 ± 0.2	13.1 ± 0.3	15.2 ± 0.3	14.8 ± 0.3	13.5 ± 0.3	13.7 ± 0.3
(c): Deep representations with channel-specific CNN modeling as here proposed							
SSE	*STD*	**8.1 ± 0.2**	**6.8 ± 0.2**	**4.8 ± 0.3**	**5.1 ± 0.3**	**7.0 ± 0.3**	**7.2 ± 0.3**
SSE	*LTD*	**8.4 ± 0.2**	**10.2 ± 0.3**	**10.7 ± 0.3**	**10.4 ± 0.3**	**8.2 ± 0.3**	**8.0 ± 0.3**
MSE	*STD*	**4.8 ± 0.2**	**4.8 ± 0.2**	**5.2 ± 0.2**	**5.1 ± 0.2**	**5.2 ± 0.2**	**5.1 ± 0.2**
MSE	*LTD*	**5.9 ± 0.3**	**7.2 ± 0.3**	**9.0 ± 0.3**	**8.7 ± 0.3**	**6.2 ± 0.3**	**6.3 ± 0.3**

Table 4. Performance (EER, in %) comparison, in task-independent conditions with $C = 19$, between EEG-based systems using representations based on: (a): hand-crafted features; (b): features learned using channel-specific CNN models, with training performed on EEG signals from a single protocol; (c): features learned using channel-specific CNN models, with training performed on EEG signals from four different protocols. A single frame lasting 5s is used as verification probe. Best results highlighted in bold.

Test Condition		Single-task enrolment	Multiple-task enrolment: Cross-task verification	Multiple-task enrolment: Within-task verification
(a): Hand-crafted representations (AR+MFCC)				
SSE	*STD*	30.6 ± 0.6	18.5 ± 0.4	18.2 ± 0.3
SSE	*LTD*	22.4 ± 0.4	20.8 ± 0.4	20.5 ± 0.4
MSE	*STD*	28.5 ± 0.5	15.7 ± 0.3	15.3 ± 0.3
MSE	*LTD*	31.5 ± 0.5	19.1 ± 0.4	18.8 ± 0.3
(b): Deep representations learned on EEG signals from a single protocol				
SSE	*STD*	19.3 ± 0.4	14.8 ± 0.3	13.2 ± 0.3
SSE	*LTD*	21.2 ± 0.4	15.6 ± 0.3	15.1 ± 0.3
MSE	*STD*	16.9 ± 0.3	12.0 ± 0.3	10.5 ± 0.3
MSE	*LTD*	19.4 ± 0.3	13.6 ± 0.3	13.1 ± 0.3
(c): Deep representations learned on EEG signals from multiple protocols				
SSE	*STD*	**14.4 ± 0.4**	**12.5 ± 0.3**	**9.8 ± 0.3**
SSE	*LTD*	**16.4 ± 0.4**	**14.0 ± 0.3**	**12.9 ± 0.3**
MSE	*STD*	**11.1 ± 0.3**	**9.7 ± 0.3**	**7.5 ± 0.2**
MSE	*LTD*	**13.5 ± 0.3**	**12.0 ± 0.3**	**10.5 ± 0.3**

fifth one, have been used to define the pairs used for training. The fourth session has been instead employed for validation purposes, creating pairs with the first three ones. For each pair of frames belonging to the same user, two additional pairs with frames from distinct subjects have been employed during siamese network training.

Stochastic gradient descend with momentum (SGDM), batches with size 128, learning rate at 0.001, and weight decay at 0.005, have been employed to train the considered siamese network using the MatConvNet deep learning framework, (Vevaldi and Lenc, 2015), with an Nvidia GeForce GTX GPU.

As already shown in (Maiorana, 2019) for the EC protocol only, the reported equal error rates (EERs) show that using EEG deep representations learned with end-to-end siamese networks guarantees recognition results better than those achievable exploiting hand-crafted features. With respect to (Maiorana, 2019), a further improvement is here obtained performing channel-specific CNN modeling. Learning different discriminative features for distinct areas of the brain allows to notably reduce error rates, with a performance percentage improvement of up to about 40%. It is therefore highly recommendable to process the available EEG channels separately when trying to learn discriminative representations for recognition purposes.

The proposed approach allows reaching quite low EERs, especially for multiple-session enrolment conditions. It is in fact possible to perform verification with EER at about 5% using EEG recognition probes lasting only 5s, taken at a time distance of one month from enrolment, for several acquisition protocols. Interestingly, thanks to the employed siamese approach and the comparison of pairs from different sessions during training, the deep representations here learned also allows to notably reduce both the performance gap between recognition rates achievable in EC conditions and using other protocols, as well as the gap between results obtained in *STD* and *LTD* conditions, with respect to the use of hand-crafted features.

4.2. Task-Independent Recognition

Several tests have been carried out to properly evaluate the feasibility of task-independent EEG-based biometric recognition. Specifically, scenarios in which users' enrolment is carried out recording EEG signals during either a single or multiple tasks have been taken into account. In the former case, verification is performed in cross-task conditions, using EEG signals recorded with protocols different than the enrolment one. The multi-protocol registration scenario has been instead analyzed performing several iterations in which, each time, four out of the six available protocols have been employed for enrolment purposes. Verification has been then performed using:

- EEG signals recorded according to the two tasks not considered during enrolment. As for the considered single-protocol enrolment condition, this cross-task scenario has been evaluated to assess the feasibility of recognizing a subject while performing mental tasks which have not been taken into account at enrolment stage;
- EEG signals recorded with the same four protocols employed for users' enrolment. This within-task scenario has been evaluated to investigate the performance achievable when performing enrolment recording signals in all the conditions which could be faced during verification.

The results reported in Table 4 have been obtained considering three different kinds of EEG models:

Table 5. Average EER (in %), in cross-task experimental conditions and using deep representations learned on EEG signals from multiple protocols, for each channel. Best channels highlighted in bold. A single frame lasting 5s is used as verification probe.

Test Condition		F_1	F_2	F_3	F_4	F_7	F_8	F_Z	C_3	C_4	C_Z	T_3	T_4	T_5	T_6	P_3	P_4	P_Z	O_1	O_2
MSE	*STD*	26.1	27.5	24.77	26.0	29.3	27.3	**21.7**	**21.0**	**22.3**	**23.2**	28.6	26.2	24.7	26.7	**21.0**	**19.7**	**21.1**	**22.4**	**22.7**
	LTD	27.8	29.0	28.2	27.3	32.2	31.6	**24.3**	**24.3**	**24.1**	**23.7**	29.1	27.5	27.2	27.1	**22.3**	**22.0**	**22.4**	**23.5**	**25.0**

- hand-crafted representations based on fusion of AR and MFCC features as in (Maiorana and Campisi, 2018);
- deep representations defined through the proposed channel-specific siamese CNN model, using EEG data acquired according to a single protocol, as in the Section 4.1;
- deep representations defined according to the channel-specific approach here proposed, exploiting EEG data recorded with four different protocols. At each iteration, the considered protocols are the same exploited in the multi-enrolment scenarios. Obviously, given the open-set testing conditions here evaluated, subjects other than those employed to estimate recognition performance have been used for network training. It is worth remarking that genuine pairs of EEG frames employed to train the proposed siamese networks have been defined taking samples related both to different recording sessions, as in Section 4.1, as well as to different acquisition protocols. The network has been thus forced to explicitly learn task-independent EEG discriminative characteristics.

From the achieved performance it is possible to observe that performing task-independent EEG-based recognition relying on hand-crafted features is really arduous. High error rates are in fact obtained even in case of multiple-protocol enrolment, with further worsening when EEG data recorded with a single protocol are used for enrolment.

Also the results obtained when learning EEG representations using a single protocol during network training are not satisfactory. It can be noticed that enrolling a subject using a protocol, and performing verification with a different one, involves a notable performance worsening with respect to the exploitation of the same task in both phases, as done for the tests summarized in Table 3. Yet, considering more than a single protocol during enrolment may improve the achievable recognition performance: considering for instance the comparison of EEG signals taken at a short time distance, the achievable EER improves from 16.9% for single-session enrolment conditions, to 12.0% using data from multiple sessions for enrolment. As expected, performing verification on EEG signals acquired performing the same tasks considered during enrolment further improves the obtained results, with an EER at 10.5% in the aforementioned conditions.

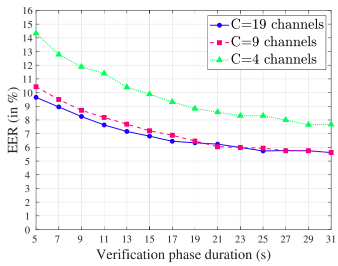
The most interesting results have been achieved when training the proposed channel-specific siamese networks with EEG signals acquired according to distinct protocols. Such approach allows notably improving the recognition performance attainable in both single-protocol and multiple-protocol enrolment conditions, being for instance possible to respectively achieve EERs at 11.1% and 9.7% in $\{MSE,STD\}$ scenario, when performing verification using EEG signals recorded with protocols other than those acquired during enrolment. The performed siamese training therefore actually allows defining EEG representations containing task-independent discriminative information. It is also worth observing that performing verification exploiting EEG signals acquired with the same protocols considered during enrolment, and resorting to representations learned on the same tasks, allows achieving the best EERs, obtaining for instance a 7.5% in the $\{MSE,STD\}$ scenario.

The obtained results suggests that, in order to guarantee the best possible recognition rates, it would be advisable to train neural networks employing EEG signals acquired in as many conditions as possible, and applying the derived representations to EEG data recorded during as many tasks as possible during users' enrolment too.

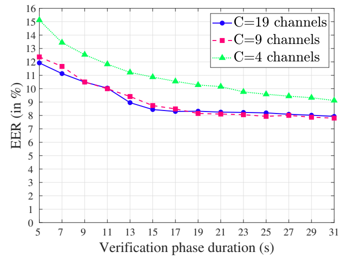
4.3. Improving Usability: Channels Reduction

Further tests have been carried out to investigate the feasibility of performing EEG-based biometric recognition while taking into account usability issues. Specifically, given that the proposed approach separately processes different EEG channels with dedicated CNNs, the possibility of reducing the employed channels while guaranteeing proper recognition rates has been evaluated. Table 5 reports, for each individual channel, the EERs achieved when comparing EEG signals at both short and long time distances, adopting deep representations learned on EEG data captured with four different protocols, using multiple enrolment sessions with each of them comprising four protocols, and performing verification with EEG signals recorded during tasks other than those employed during enrolment, as in the cross-task verification scenario of Table 4(c). The obtained results are significantly consistent with the analysis conducted in (Maiorana and Campisi, 2018). The best recognition rates have been here achieved exploiting a subset of four channels comprising the $\{C_Z, P_Z, P_3, P_4\}$ positions. Relevant discriminative capabilities are also exposed by the channels in the $\{F_Z,C_3,C_4,O_1,O_1\}$ subset. The major difference, compared to the results in (Maiorana and Campisi, 2018) where hand-crafted representations have been considered, lies in the greater discriminative capability observed in channels C_3 and C_4, with respect to F_3 and F_4, preferred in (Maiorana and Campisi, 2018). Channels F_7 and F_8 instead remain among the least preferable.

The behaviors achievable for verification phases with increasing duration, fusing the scores computed for each frame, are then shown in Figure 3. The reported plots refer to cross-task verification scenarios with enrolment data captured at multiple sessions (*MSE*), each time considering more than a single protocol, deep representations learned on EEG data captured with four different protocols, and both *STD* and *LTD* between enrolment and verification. The obtained results show that using the $C = 9$ best channels guarantees the same recognition



(a)



(b)

Fig. 3. Temporal behavior of the EER achievable in cross-task comparisons with different number of employed channels (a): *STD*, (b): *LTD*.

results achievable using all the $C = 19$ available ones. Further reducing the employed signals to only $C = 4$ channels involve a performance worsening, with EER in cross-task conditions below 10% achievable when using probe EEG samples lasting at least 15s in *STD* comparisons, and exploiting at least 21s of EEG signals under *LTD* scenarios.

5. Conclusions

The feasibility of performing task-independent EEG-based biometric recognition has been evaluated in this paper. A multi-session and multi-protocol database comprising EEG recordings from 45 subjects has been employed to train channel-specific CNNs, and test the discriminative capability of the derived templates on test datasets disjoint from the training ones. The employed siamese training strategy has allowed to learn EEG representations which can be therefore employed to perform EEG-based biometric verification under cross-task conditions, comparing EEG recordings taken at time distances even greater than one year. An analysis of the recognition performance achievable when reducing the number of the employed electrodes, with the aim of improving the usability of EEG-based biometric systems, has been also conducted.

The performed tests have highlighted the importance of employing deep learning approaches to learn EEG representations invariant with respect to the performed mental tasks. Further studies, focusing for instance on generative adversarial networks or autoencoders to model EEG signals, are however needed to improve the achievable recognition rates.

References

Bianco, S., Napoletano, P., 2019. Biometric recognition using multimodal physiological signals. IEEE Access 7, 83581–83588.

Campisi, P., La Rocca, D., 2014. Brain waves for automatic biometric-based user recognition. IEEE Transactions on Information Forensics and Security 9, 782–800.

Chen, Y., Atnafu, A., Schlattner, I., Weldtsadik, W.T., Roh, M.C., Kim, H., Lee, S.W., Blankertz, B., Fazli, S., 2016. A high-security EEG-based login system with RSVP stimuli and dry electrodes. IEEE Transactions on Information Forensics and Security 11.

da Silva Luz, E., Moreira, G., Oliveira, L., Schwartz, W., Menotti, D., 2018. Learning deep off-the-person heart biometrics representations. IEEE Transactions on Information Forensics and Security 13, 1258–1270.

Del Pozo-Banos, M., Travieso, C., Alonso, J., John, A., 2018. Evidence of a task-independent neural signature in the spectral shape of the electroencephalogram. International Journal of Neural Systems 28.

Fraschini, M., Pani, S., Didaci, L., Marcialis, G., 2019. Robustness of functional connectivity metrics for EEG-based personal identification over task-induced intra-class and inter-class variations. Pattern Recognition Letters 125, 49–54.

Friston, K.J., 2011. Functional and effective connectivity: A review. Brain Connectivity 1, 13–36.

Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P., Mark, R., Mietus, J., Moody, G., Peng, C.K., Stanley, H., 2000. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation 101, 215–220.

Gui, Q., Ruiz-Blondet, M., Laszlo, S., Jin, Z., 2019. A survey on brain biometrics. ACM Computing Surveys 51, 1–38.

Kong, X., Kong, W., Fan, Q., Zhao, Q., Cichocki, A., 2018. Task-independent EEG identification via low-rank matrix decomposition, in: IEEE International Conference on Bioinformatics and Biomedicine (BIBM).

Kumar, M., Saranya, M., Narayanan, S., Sur, M., Murthy, H., 2019. Subspace techniques for task-independent EEG person identification, in: IEEE Conference of the Engineering in Medicine and Biology Society (EMBC).

Ma, L., Minett, J., Blu, T., Wang, W.Y., 2015. Resting state EEG-based biometrics for individual identification using convolutional neural networks, in: IEEE EMBC.

Maiorana, E., 2019. EEG-based biometric verification using siamese CNNs, in: International Workshop on Recent Advances in Digital Security: Biometrics and Forensics.

Maiorana, E., 2020. Deep learning for EEG-based biometric recognition. Neurocomputing 410, 374–386.

Maiorana, E., Campisi, P., 2018. Longitudinal evaluation of EEG-based biometric recognition. IEEE transactions on Information Forensics and Security 13, 1123–1138.

Maiorana, E., La Rocca, D., Campisi, P., 2016a. On the permanence of EEG signals for biometric recognition. IEEE Transactions on Information Forensics and Security 11, 163–175.

Maiorana, E., Solé-Casals, J., Campisi, P., 2016b. EEG signal preprocessing for biometric recognition. Machine Vision and Applications 27, 1351–1360.

Niedermeyer, E., Da Silva, F., 2005. Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Doody's all reviewed collection, Lippincott Williams & Wilkins.

Ozdenizci, O., Wang, Y., Koike-Akino, T., Erdogmus, D., 2019. Adversarial deep learning in EEG biometrics, pp. 710–714.

Pinegger, A., Wriessnegger, S., Faller, J., Muller-Putz, G., 2016. Evaluation of different EEG acquisition systems concerning their suitability for building a brain computer interface: Case studies. Frontiers in Neuroscience 10, 1–11.

Pinto, J., Cardoso, J., Lourenco, A., 2018. Evolution, current challenges, and future possibilities in ECG biometrics. IEEE Access 6, 34746–34776.

Revett, K., 2012. Cognitive biometrics: a novel approach to person authentication. Int. Jour. Cognitive Biometrics 1, 1–9.

Ruiz-Blondet, M., Jin, Z., Laszlo, S., 2016. CEREBRE: A novel method for very high accuracy event-related potential biometric identification. IEEE Transactions on Information Forensics and Security 11, 1618–1629.

Stassen, H.H., 1980. Computerized recognition of persons by EEG spectral patterns. Electroencephalography and Clin. Neurophysiology 49, 190–194.

Vevaldi, A., Lenc, K., 2015. MatConvNet - convolutional neural networks for Matlab, in: ACM International Conference on Multimedia.

Vinothkumar, D., Kumar, M., Kumar, A., Gupta, H., Saranya, M., Sur, M., Murthy, H., 2018. Task-independent EEG based subject identification using auditory stimulus, in: Workshop on Speech, Music and Mind.

Wang, M., El-Fiqi, H., Hu, J., Abbass, H., 2019. Convolutional neural networks using dynamic functional connectivity for EEG-based person identification in diverse human states. IEEE Transactions on Information Forensics and Security 14, 3259–3272.

Yang, S., Deravi, F., 2017. On the usability of electroencephalographic signals for biometric recognition: A survey. IEEE Transactions on Human-Machine Systems 47, 958–969.