MOTOR IMAGERY FOR EEG BIOMETRICS USING CONVOLUTIONAL NEURAL NETWORK

Rig Das, Emanuele Maiorana, Patrizio Campisi

Section of Applied Electronics, Department of Engineering, Roma Tre University
Via Vito Volterra 62, 00146 Roma, Italy
{rig.das, emanuele.maiorana, patrizio.campisi}@uniroma3.it

ABSTRACT

This paper deals with electroencephalography (EEG)-based biometric identification, using a motor imagery task, specifically performing imaginary arms and legs movements. Deep learning methods such as convolutional neural network (CNN) is used for automatic discriminative feature extraction and person identification. An extensive set of experimental tests, performed on a large database comprising EEG data collected from 40 subjects over two different sessions taken at a week distance, shows the existence of repeatable discriminative characteristics in individuals' brain signals.

***Index Terms*—** Electroencephalography, Motor imagery, Convolutional neural network.

1. INTRODUCTION

The availability of automatic biometric recognition systems, able to measure unique physical and behavioral characteristics of an individual to verify his identity, is of utterly importance for many practical applications. Such need has fostered the current widespread usage of several biometric traits such as fingerprint, iris, and face just to cite a few. However, these traits are often vulnerable to relevant issues, implying the possibility of being forged or used for unethical purposes [1]. The design of innovative biometric frameworks, based on either protected architectures [2] or novel secure identifiers [3], is therefore a highly prominent topic. In this regard, brain signals have recently attracted the attention of the scientific community, on the basis of the postulated assumption that our brain possesses peculiar subject-specific properties. It is in fact worth remarking that brain activity cannot be covertly sensed without subjects' cooperation, is much harder to be forged than traditional biometrics, and inherently guarantees liveness detection [4, 5]. Among the different modalities that can be used to sense brain's activity for people recognition, electroencephalography (EEG) has received most of researchers' interest, since it permits collecting brain information using portable and relatively inexpensive devices, a notable advantage to raise the adoption of such trait in practical biometric systems [6]. EEG signals are generated by the synchronous firing of specific spatially-aligned neurons of the cortex, i.e., pyramidal neurons, and can be measured as voltage differences by metal electrodes placed on the head scalp surface. The characteristics of the collected data depend on the specific acquisition protocol employed to elicit a given subject's behavior. Different task-related responses, typically expressed in form of small time-locked changes in the electrical activity of the brain, can in fact be obtained using various stimulation paradigms involving sensory or cognitive audio/video stimuli.

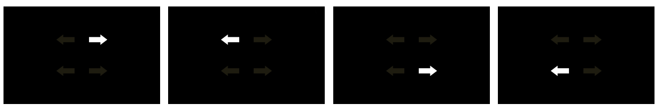
The present research work specifically exploits EEG signals elicited during motor imagery (MI) tasks for automatic biometric identification. In more detail, MI refers to a specific cognitive process during which a subject imagines to perform arms or legs movements without actually performing them. An imagined movement requires a conscious activation of specific brain regions involved in movement preparation and execution, typically accompanied by a voluntary prevention of the actual movement [7]. It propagates spontaneous time-locked brain signals from the visual cortex to the subjects' scalp, where they can be recorded through EEG. Previous researches show that motor imagery is a suitable technique for the design of brain computer interfaces (BCI), helping disabled people communicate and control devices [8]. It is however worth remarking that eliciting MI-based EEG signals requires high alertness and willingness of the involved subject, since several imaginary movements have to be collected to provide information useful for recognition purposes. Moreover, recorded data can also get contaminated by eye-blinking artifacts and noise. Hence, extracting discriminative features from the acquired data is a not-trivial task, difficult to be performed through manual or simplistic approaches. In this contribution, we exploit the properties of deep learning methods such as convolutional neural networks (CNN) to automatically extract subject-specific features, and classify them for people identification purposes.

2. MOTOR IMAGERY FOR EEG BIOMETRICS

A brief summary of the state-of-the-art works using MI-elicited EEG signals for biometric recognition is provided in Table 1. In [8], authors have exploited MI EEG data captured through six channels from three subjects, elicited to perform hands, feet and tongue movements during a single acquisition session. Auto-regressive (AR) and moving average (ARMA) coefficients have been used for feature extraction, while classification has been performed through multi-layer

Table 1. Overview of state-of-the-art EEG based biometric systems based on the use of MI tasks.

Paper	Users	Channels	Type of Task	Features	Classifier	Performance	Sessions
Hu [8]	3	6 (C_3,C_4,P_3,P_4,O_1,O_2)	Left & right hand, foot, tongue	AR/ARMA coefficient	Multi-layer BP-NN	CRR=$81.9 - 83.9\%$	1
Xiao *et al.* [9]	3	60	Left & right hand, foot, tongue	Fisher distance	BP-NN	CRR=80%	1
Marcel *et al.* [10]	9	8 (centro-parietal)	Left & right hand	Power spectral density	GMM	HTER=7.1%	3
Tsuru *et al.* [11]	9	4 (F_z,C_3,C_4,P_z)	Left & right hand, foot, tongue	Cepstral Values	Mahalanobis	EER=0.17%	2
Yang *et al.* [12]	108	9 (AF_3,AF_z,AF_4,C_1 C_z,C_2,O_1,O_z,O_2)	Left/right fist	Wavelet packet decomposition coefficient	LDA	CRR=94.72%	1
			Both fists & feet			CRR=93.1%	



(a) Right-Arm (b) Left-Arm (c) Right-Leg (d) Left-Leg

Fig. 1. Images of “arrows” for imaginary limbs movement

back propagation neural network (BP-NN). Identification accuracies ranging from 81.9% to 83.9% have been achieved. In [9] authors have achieved 80% correct recognition rate (CRR) using Fisher distance as characteristic feature and BP-NN as classifier, for the same number of subjects, session and imaginary tasks as the previous case, while 60 EEG channels have been instead exploited. In [10], EEG data elicited through imaginary left and right hand movements have been acquired from nine subjects in three different sessions, using eight central and parietal channels to extract power spectral density (PSD) features. A half total error rate (HTER) of 7.1% has been achieved using Gaussian mixture models (GMMs) as classifier. EEG data from nine subjects have been collected in two different days in [11], using four channels and according to four standard MI tasks. Cepstral values have been used as features to achieve 94.72% accuracy with Mahalanobis distance as classifier. In [12], EEG signals have been acquired through nine channels in a single acquisition session from 108 subjects performing two protocols, based on either left/right fist or both fists & feet’s imaginary movements. A 94.72$\%$ accuracy has been achieved for the former protocol, while 93.1$\%$ has been obtained for the latter, using wavelet packet decomposition for feature extraction and linear discriminant analysis (LDA) as classifier. It can be observed that the studies carried out so far report recognition performance evaluated over either EEG data collected from a very small number of subjects, or data recorded during a single acquisition session, which cannot provide any convincing evidence for considering EEG signals as a stable biometric identifier. In fact, under such conditions it is hard to state whether the reported recognition performance depend only on the characteristics of each subject’s neural activity, or also on session-specific exogenous conditions, such as the capacitative coupling of electrodes and cables with lights or computer, induction loops created between the employed equipment and the body, power supply artifacts, and so on. Considering the limits of the contributions proposed so far, the current work investigates the stability and invariability of EEG signals over a 40 subjects’ database, acquired during two distinct sessions which are separated by a week of time.

3. USED PROTOCOL & TEMPLATE GENERATION

3.1. “Motor Imagery” Protocol

The MI protocol is designed to collect EEG signals corresponding to a series of right-arm, left-arm, right-leg and left-leg imaginary movements. Four different images, depicted in Fig.1, are employed as stimuli for eliciting the desired responses and are sequentially shown in random order on a LCD monitor, while a subject sitting on a relaxing chair with armrests is asked to perform the imaginary movement associated to the image currently displayed. The performed task should trigger high-amplitude brain signals, which are then acquired through N EEG channels for further processing. Each of the employed stimuli is randomly selected and displayed for 3s during 50 occurrences. An empty black screen lasting 1.5s is shown in between every two consecutive images, and a 6s rest is allowed each time the whole set of four stimuli has been presented for 5 times.

3.2. Template Generation

The N acquired EEG signals are pre-processed using a common average referencing (CAR) spatial filter, in order to reduce the artifacts related to a possible unsuitable reference. The obtained data are then spectral filtered into the $\mu = [8 : 12]$ Hz sub-band [13], supposed to contain the most-significant information available in MI EEG signals [4]. A down-sampling to 128 Hz is performed, from the employed 256 Hz rate, to reduce the computational complexity. EEG signals are then normalized to generate zero-mean data with unit variance, and de-trended by subtracting their best-fit line. Since MI EEG potentials’ amplitude is usually significantly low in comparison with the overall behavior of the observed fluctuations, further processing involving signal averaging across multiple occurrences of EEG responses to the same stimulus has to be performed, before providing the available data to the CNN. In more detail, the responses to each event are extracted from the recorded data flow and organized as epochs lasting Δ_t s after the presentation of a specific

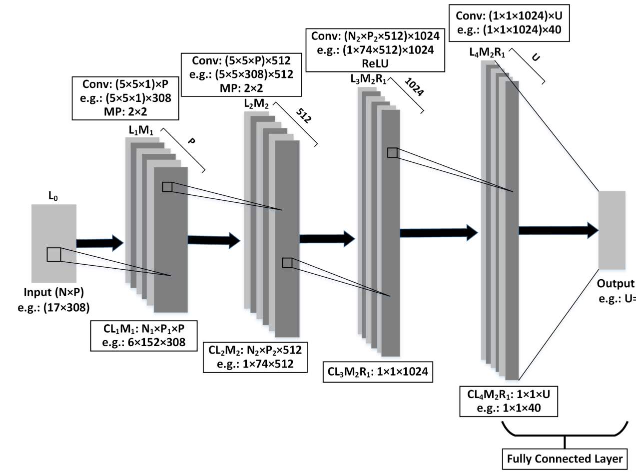


Fig. 2. Employed CNN Architecture.

stimulus. Given a N-channel EEG signal collected through the proposed protocol, mean behaviors across $R = 40$ consecutive responses to the same stimulus are evaluated from the available epochs to filter out the undesired noise, with a maximum of $T = 50$ averaged templates which can be therefore generated for each imaginary movement. Having observed in the performed experimental tests that it is difficult to achieve high identification accuracy exploiting a single MI task among the considered ones, a feature-level fusion of the available information is performed by horizontally concatenating the computed average EEG responses to the four considered stimuli. This generates a template representation for the EEG data of a specific user in the form of a T matrix having size $[N \times P]$, where $P = 4 \cdot 128 \cdot \Delta_t$ is the number of samples obtained concatenating four signals, each lasting Δ_t s and captured at 128 Hz.

4. NETWORK TOPOLOGY & TRAINING

A deep learning method leveraging on CNNs is proposed in this paper to perform both automatic discriminative feature extraction and person identification in a biometric identification systems exploiting MI EEG signals. A CNN is a multi-layer perceptron (MLP) network with a special topology containing more than one hidden layer [14, 15]. CNN is in fact advantageous when dealing with input data having a specific, possibly unknown, inner structure, with the purpose of discovering invariant distinctive features from them [16]. Such approach can be therefore suitable for dealing with EEG data, which typically exhibit substantial variability over time and individuals, making it hard to cope with them through classical local-kernel-based architectures.

4.1. Network Topology

The adopted CNN network topology is shown in Fig.2. It is designed to have 4 conv layers, 2 max-pooling, 1 rectified linear unit (ReLU), and a softmax-loss layer. Specifically, the templates generated as described in Section 3.2 are provided to the first hidden layer, where a set of very low-level features of the signals are extracted. In the subsequent conv layers the network built on such low-level features, and eventually high-level features are generated in the fully connected layer. The CNN is implemented using MatConvNet-1.0-beta24 [17] packages, in a system configuration comprising 64 GB RAM, a 12 GB TITAN X (Pascal) graphics card, an i7, $3.40GHz$ processor, and Windows 10 operating system. The detailed network topology is described as follows:

- L_0: the input layer has an input data size of $[N \times P]$, as described in Section 3.2;
- L_1M_1: first hidden layer, composed of P conv filters of size $[W \times W \times 1]$, and a max-pooling (MP) layer of size $[B \times B]$. This layer transforms the input data into a representation having size $CL_1M_1 = [N_1 \times P_1 \times P] = [\lfloor \frac{N-W+1}{B} \rfloor \times \lfloor \frac{P-W+1}{B} \rfloor \times P]$, after convolving and down-sampling. $W = 5$ and $B = 2$ are here used;
- L_2M_2: second hidden layer, composed of 512 conv filters of size $[W \times W \times P]$ and a max-pooling layer of size $[B \times B]$. This layer transforms the first hidden layer's data into a representation having size $CL_2M_2 = [N_2 \times P_2 \times 512] = [\lfloor \frac{N_1-W+1}{B} \rfloor \times \lfloor \frac{P_1-W+1}{B} \rfloor \times 512]$;
- $L_3M_2R_1$: the third hidden layer is composed of 1024 conv filters of size $[N_2 \times P_2 \times 512]$ and a ReLU layer, whose purpose is to introduce non-linearity into the system. This layer changes the previous layer's output into a $CL_3M_2R_1 = [1 \times 1 \times 1024]$ feature map;
- $L_4M_2R_1$: the output layer is produced by convolving the previous layer's activation map with U conv filters of size $[1 \times 1 \times 1024]$, being U the overall number of subjects enrolled in the considered biometric identification system. This layer has only one map of U neurons representing the U classes/subjects, and is fully connected with $L_3M_2R_1$. Softmax-loss function is here used as a loss function for back-propagation.

5. EXPERIMENTAL SETUP

5.1. Employed EEG Database

EEG data from 50 healthy subjects, whose age ranges from 20 to 35 years, have been collected according to the MI protocol described in Section 3.1 during two distinct recording sessions, indicated as $S1$ and $S2$, temporally separated by a week period. The employed EEG data acquisition system is a Galileo BE Light amplifier with 19 electrodes, placed on the subjects' scalp according to the 10-20 international system [18]. Out of 19 available electrodes, we consider for experimental tests only N= 17 channels, excluding the two frontal ones, i.e., F_{p1} and F_{p2}, as most relevant EEG potentials for motor imagery protocol are present in the central and parietal regions of the brain [4].

5.1.1. Subject Selection

EEG responses to a MI protocol requires the voluntary activation of specific brain regions by the involved subjects. This implies that, although subjects are asked to concentrate during

EEG acquisition, some may not properly perform the required imaginary tasks. A preliminary analysis on the available data, to check whether the signals acquired from the 50 considered subjects are suitable for further processing, has then to be performed. In order to do so, for each acquired subject, the mean band-power measured over all the 50 occurrences of each specific task is evaluated for all the N available channels, and then averaged. If the difference in band-power estimated for the same stimulus between the two performed acquisition sessions is higher than a predefined threshold, the considered subject is excluded from training and testing datasets in further processing. In fact, a too high difference in band-power would imply the subject could have not properly performed the requested task, at least in one of the acquisition sessions. A careful investigation of the data acquired from all the 50 considered subjects has highlighted that 7 of them behaved too differently during all the four performed MI tasks, while other 3 performed badly for at least 2 or 3 tasks. Such subjects have been therefore excluded from further processing, since their data have been assumed to be unreliable. A total of $U = 40$ subjects have been therefore considered.

5.2. CNN Training

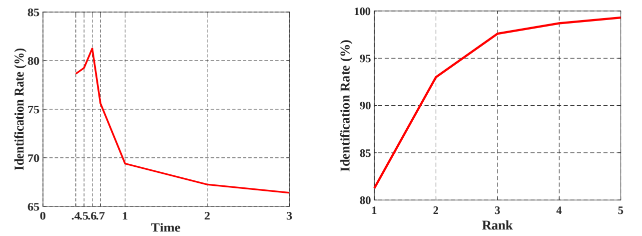
The employed CNN is trained using EEG data collected from $U = 40$ enrolled subjects during $S1$ session. The training dataset is an ensemble of the signals from all the considered subjects, represented through a matrix having size $[N \times P \times U \cdot T]$. 90% of each subject's data is used for training purposes, while and rest 10% is employed for validating the build network. The learning rate of the CNN network is set at 0.0001 with a batch size of 5, so that the loss can be minimized with higher precision along with the execution of every iteration. For our experiment we have investigated 50, 100 200 and 300 iterations, and found that a number of 200 is enough to achieve optimal accuracy.

5.3. Identification

The testing templates used for the identification stage are generated from EEG signals captured during session $S2$. Training and testing datasets are therefore completely disjoint. For every testing sample of size $N \times P$, the trained CNN network returns probability values corresponding to all the $U = 40$ classes/subjects. The maximum probability value identifies the subject with which the testing sample is more similar.

6. RESULTS & DISCUSSION

Several time intervals $\Delta_t = [0.4, 0.5, 0.6, 0.7, 1, 2, 3]\ s$, having length up to the duration of the stimuli employed in the elicitation protocol, are considered to check which duration of the EEG response to MI stimuli is the best suited for biometric identification. Being the size of the employed templates dependent on the exploited time interval, the proposed network is *ad-hoc* designed for each considered scenario, as described in Section 4.1. Fig.3(a) shows the rank-1 identification rate for different values of Δ_t. Results show that the best performance can be obtained considering EEG signals lasting



(a) (b)

Fig. 3. Identification rate for (a) Time interval selection and (b) Cumulative Match Curve (CMC) for N=17.

$600\ ms$, while worse identification rates are obtained taking into account more information. This implies that a practical MI protocol can be designed to be much shorter than the one we evaluate, given that users can be assumed to get tired or loose concentration after few hundreds of milliseconds following the stimuli presentation. The value $\Delta_t = 600\ ms$ is therefore kept as duration of the considered MI responses for further testing. The performance obtained as rank-wise identification rates is shown in Fig.3(b). Rank-1 and rank-2 results are respectively set at 81.25% and 93% accuracy, showing a significant increase in performance at rank-2 over the considered database with $U = 40$ subjects. The achieved accuracy then reaches 99.3% for rank-5 identification. In comparison with the state-of-the-art methods our EEG data are collected from 2 distinct session which are separated by a week and the CNN network training and testing are performed on two distinct session. Therefore the obtained results confirm the assumption that EEG data possess permanent discriminative characteristics, therefore providing valuable information encouraging the adoption of brain signals for futuristic biometric identification systems.

7. CONCLUSIONS

A biometric identification system based on EEG signals elicited through a MI protocol has been proposed in this paper. A CNN has been designed to automatically perform feature extraction and classification from the collected data, allowing to perform user identification even in case EEG data from different acquisition sessions are used for training and testing purposes. EEG biometrics anyhow provides high levels of confidentiality, universality and security, to improve the performance achievable in multi-biometrics approaches. It is also worth remarking that EEG-based biometric systems might be useful for physically-impaired people who may be unable to use conventional biometric recognition systems, and it may also be used in law enforcement or defense-related recognition systems demanding high security.

Acknowledgment

We gratefully acknowledge the support of NVIDIA® Corporation, providing the Titan X™ GPU used for this research.

8. REFERENCES

[1] A. K. Jain, K. Nandakumar, and A. Ross, "50 years of biometric research: Accomplishments, challenges, and opportunities," *Pattern Recognition Letters*, vol. 79, pp. 80–105, 2016.

[2] P. Campisi, Ed., *Security and Privacy in Biometrics*, Springer, 2013.

[3] K. Revett, "Cognitive biometrics: a novel approach to person authentication," *Int. Jour. Cognitive Biometrics*, vol. 1, no. 1, pp. 1–9, 2012.

[4] P. Campisi and D. La Rocca, "Brain waves for automatic biometric-based user recognition," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 5, pp. 782–800, May 2014.

[5] E. Maiorana and P. Campisi, "Longitudinal evaluation of eeg-based biometric recognition," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1123–1138, May 2018.

[6] J. Klonovs, C.K. Petersen, H. Olesen, and A. Hammershoj, "ID proof on the go: Development of a mobile EEG-based biometric authentication system," *IEEE Vehicular Technology Magazine*, vol. 8, no. 1, pp. 81–89, 2013.

[7] T. Mulder, "Motor imagery and action observation: cognitive tools for rehabilitation," *Springer Journal of Neural Transmission*, vol. 114, no. 10, pp. 1265–1278, 2007.

[8] J.F. Hu, "New biometric approach based on motor imagery EEG signals," in *2009 International Conference on Future BioMedical Information Engineering (FBIE)*, Dec 2009, pp. 94–97.

[9] D. Xiao and J. Hu, "Identification of motor imagery EEG signal," in *2010 International Conference on Biomedical Engineering and Computer Science*, April 2010, pp. 1–4.

[10] S. Marcel and J. D. R. Millan, "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 743–752, April 2007.

[11] K. Tsuru and G. Pfurtsheller, "Brainwave biometrics:a new feature extraction approach with the cepstral analysis method," *Transactions of Japanese Society for Medical and Biological Engineering*, vol. 50, no. 1, pp. 162–167, 2012.

[12] S. Yang, F. Deravi, and S. Hoque, "Task sensitivity in EEG biometric recognition," *Pattern Analysis and Applications*, pp. 1–13, 2016.

[13] B. J. Edelman, B. Baxter, and B. He, "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 1, pp. 4–14, Jan 2016.

[14] H. Cecotti and A. Graser, "Convolutional neural networks for P300 detection with application to brain-computer interfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, pp. 433–445, March 2011.

[15] D. Yu and L. Deng, "Deep learning and its applications to signal and information processing," *IEEE Signal Processing Magazine*, vol. 28, no. 1, pp. 145–154, 2011.

[16] L. Ma, J. W. Minett, T. Blu, and W. S. Y. Wang, "Resting state EEG-based biometrics for individual identification using convolutional neural networks," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2015, pp. 2848–2851.

[17] A. Vedaldi and K. Lenc, "Matconvnet – convolutional neural networks for MATLAB," in *Proceeding of the ACM Int. Conf. on Multimedia*, 2015.

[18] J. Malmivuo and R. Plonsey, *Bioelectromagnetism : principles and applications of bioelectric and biomagnetic fields*, Oxford University Press, New York, 1995.