

Visually Evoked Potential 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

Email: {rig.das, emanuele.maiorana, patrizio.campisi}@uniroma3.it

Abstract—In this paper we investigate the performance of electroencephalographic (EEG) signals, elicited by means of visual stimuli, for biometric identification. A deep learning method such as convolutional neural network (CNN), is used for automatic discriminative feature extraction and individual identification. Experiments are performed on a longitudinal database comprising of EEG data acquired from 40 subjects over two distinct sessions separated by a week time. The experimental results testify the existence of repeatable discriminative characteristics in individuals' EEG signals.

Index Terms—Electroencephalography, Visually evoked potential, Convolutional neural network

I. INTRODUCTION

Electroencephalographic (EEG) signals have been theorized to possess distinctive characteristics as biometric identifiers in [1]. Recently EEG biometrics has attracted again the interest of researchers due to some convenient properties such as their confidentiality and ability to guarantee high security [2]. However, some restraints are also associated with the practice of EEG-based biometric recognition. For instance, EEG signals are highly sensitive to both endogenous and exogenous noises during acquisition, typically resulting in the presence of artifacts in the recorded data. Hence, it is difficult to perform a proper feature extraction and selection for a EEG-based biometric identification system. In this regard, several machine learning techniques, based on models such as neural network (NN) [3], hidden Markov model (HMM) [4], or support vector machine (SVM) [5], have been already proposed for people recognition using different kinds of brain signals. Back-propagation-based neural networks such as convolutional neural network (CNN) represent another interesting model which could be proficiently applied for EEG-based biometric applications. In fact, one of the most useful properties of CNN relies on the distribution of its neurons' weights once the network is trained. The convolution between these neurons and receptive fields/kernels generates a distinguishing characteristics about the type of high-level features that are there to be detected [6]. In this paper we exploit the capabilities of CNN for biometric people identification using EEG signals elicited through protocols generating visually evoked potentials (VEPs).

This specific type of brain signal is based on the fact that a visual stimulus generates spontaneous time-locked EEG responses from the visual cortex, which can be recorded and used for user recognition purposes [7]. In more detail, in

this paper we investigate a “geometric” protocol [7], where different visual stimuli are presented to the considered subjects in terms of geometric shapes, in order to generate VEP responses for individuals' biometric identification. The adopted experimental protocol include the presentation of sequences of both target and non-target stimuli images to the subjects. We perform tests on EEG recordings collected from 40 subjects during two distinct sessions, spanned over a period of one week, in order to inspect the achievable stability and identification performance across time. The accuracy achieved by our biometric identification system reestablishes our claim on the permanence of EEG signals made in [7], where it has been shown that VEP responses can be efficiently used as stable biometric identifiers across different acquisition sessions.

This paper is organized as follows: Section II provides a brief review on the state-of-the-art VEP-based biometric recognition. Section III provides the detailed description of our data acquisition protocol and CNN topology, while Section IV discusses about the employed EEG-based biometric identification system. Section V presents the obtained experimental results, and conclusions are eventually drawn in Section VI.

II. STATE OF THE ART: VEP-BASED EEG BIOMETRICS

A brief synopsis of the state-of-the-art works on the use of visual-stimuli-elicited EEG signals for biometric identification is presented in this section. This approach, for individual identification, has been first proposed in [8], where VEP signals have been recorded from 20 subjects by presenting black and white images of common objects, using 61 channels and exploiting the gamma ($[30 : 40]Hz$) band, with spectral power ratio as features. A back-propagation neural network (BPNN) has been used to identify individuals with 99.6% accuracy while performing ANOVA tests on each individual channel. In [9] EEG responses have been collected from 5 different subjects during 5 sessions on the same day. In a particular session, a sequence of 9 images has been randomly shown for 20 times to each subject, while asking him to focus on one or more pre-selected target images and ignore the rest. Principal component analysis (PCA) has been applied on the obtained time sequences for feature extraction, and linear discriminant analysis (LDA) used for classification. A performance accuracy of 97.6% has been achieved by considering only one channel for both target and non-target stimuli. The significance of irrelevant stimuli has been studied

TABLE I
OVERVIEW OF STATE-OF-THE-ART FOR VISUAL STIMULI ELICITED EEG BIOMETRIC SYSTEMS.

Paper	Users	Ch.s	Protocol	Type of Stimuli	Features	Classifier	Performance	Sessions
Palaniappan [8]	20	61	VEP	snodgrass & vanderwart pictures	spectral power ratio	BP NN	CRR=99.6%	1
Touyama [9]	5	1 (Cz)	VEP/ERP	target and non-target images	PCA	LDA	CRR=97.6%	5 (same day)
Gupta et al. [10]	8	8	VEP/ERP	rapid serial visual paradigm	P300	LDA	CRR=97.0%	1
Das et al. [11]	20	20	VEP	rapid visual categorization task	LDA related features	KNN	CRR=94.0%	1
Yeom et al. [12]	10	8	VEP/ERP	self and non-self face images	Adaptive discriminative feature	Non-Linear SVM	CRR=86.1%	2 (different days)
Armstrong et al. [13]	15	1	ERP	text reading	ERP signal	Correlation	CRR=89.0%	2 (1 week)
	8						CRR=93.0%	2 (over 6 months)

in [10] using rapid serial visual paradigm (RSVP) stimuli on 8 different subjects. EEG signals elicited from 8 channels have been acquired in a solo session, and P300 waves used as features. A threefold cross-validation using Bayesian LDA has been performed to obtain a maximum correct recognition rate (CRR) of 97%. In [11] VEP data from 20 subjects have been collected by exhibiting face and car images for 40ms each. SVM and LDA have been applied to discriminate individuals. A 94% classification accuracy has been achieved by selecting the best performing post-stimulus set, and using a k -nearest-neighbors (KNN)- based classification technique. It is worth specifying that all the above mentioned works have considered EEG data acquired on a single day to achieve high performance accuracy.

Conversely, in [12] EEG signals have been collected from 10 subjects during 2 separate sessions on different days, using a random sequence of self-face and other's-face images as visual stimuli. Each performed session has included 2 distinct runs, each comprising 50 trials, where in each trial a total of 20 images (10 self-face and 10 other's-face images) has been presented. A total of 180 trials has been selected for training with the remaining 20 used for testing, therefore mixing data from the two available sessions for enrolment purposes. An adaptive discriminant feature method has been used for extracting features, and non-linear-SVM for classification purposes, achieving a CRR of 86.1%. Two different schemes have been instead considered in [13]: first, EEG signals have been collected from 15 subjects at an inter-session temporal distance of one week. Then, only 8 subjects' signals have been recorded at a time span of 6 months. CRRs at 89.0% and 93.0% have been achieved by considering event related potentials (ERPs) as features, and signal correlation as classifier. As can be noticed, it is worth remarking that works performing tests on EEG data collected during acquisition sessions spanning different days typically report recognition performance much lower than those obtained exploiting EEG signals recorded during a single acquisition session.

Table I provides a summary of the aforementioned articles. It is also worth remarking that the performance so far achieved by the approaches already proposed have been obtained over relatively small databases, comprising at most 20 users when single-session datasets are considered, and 15 subjects for tests with multiple-session data, being therefore hard to derive reliable evidence regarding the suitability of

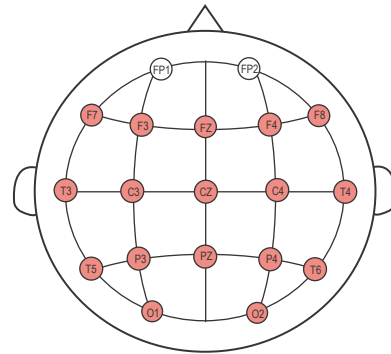


Fig. 1. Electrodes montage: $M = 17$ Ch.s.

VEP signals as biometric identifier from current literature. Considering such limits of the contributions so far presented, the present work investigates the stability and invariability of EEG signals across different acquisition sessions for the purpose of biometric recognition.

III. EMPLOYED PROTOCOL & NETWORK TOPOLOGY

According to the “geometric” visual stimulation protocol we have considered, EEG signals are elicited and acquired by presenting sequences of target and non-target stimuli. The occurrence of non-target stimuli is significantly higher than that of the target ones. The following subsections contain detailed descriptions of the employed EEG acquisition protocol, database, and CNN network topology.

A. “Geometric” Protocol

In this protocol, 8 different geometric shapes are considered as visual stimuli. Among them, the circle is the target stimulus, with the others (triangle, rectangle, square, pentagon, hexagon, octagon and diamond) treated as non-target [7]. A sequence of these images is displayed on a LCD monitor during each recording, with the observer's task being to concentrate on the occurrences of the target stimuli, while ignoring the others. A VEP is automatically elicited from the subject's brain when either target or non-target stimuli are displayed [7]. Each geometric shape is selected randomly for a total of 60 occurrences, each time displayed for 250 ms with a following empty black screen lasting 450 ms. The recording sessions therefore last a total of 5 m and 36 s for each subject.

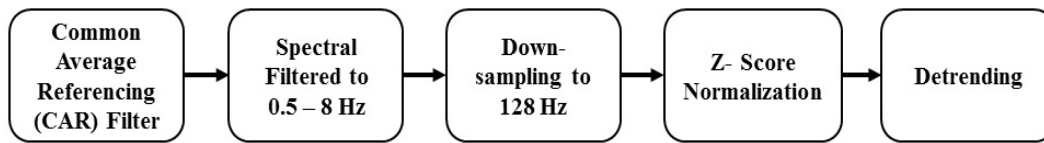


Fig. 2. Preprocessing Steps

B. Dataset & Preprocessing

A Galileo BE Light amplifier, with 19 electrodes/channels placed on the subjects' scalp according to the 10-20 international system [14], has been used for our EEG data acquisition. Specifically, we have acquired EEG data from 40 different subjects, whose age ranges from 20 to 35 years with an average of 25, by considering $M = 17$ channels, excluding the two frontal F_{P1} and F_{P2} ones from the standard 10 – 20 montage, as most relevant EEG potentials are typically in the central and occipital regions [7]. Figure 1 shows the selected 17 channels.

The data have been acquired in two distinct sessions, namely $S1$ used as training dataset, and $S2$ employed as testing dataset, with $S2$ separated by a week from $S1$. EEG signals are pre-processed using a common average referencing (CAR) filter in order to reduce the artifacts that are related to unsuitable reference. Subsequently we perform a frequency filtering on the acquired EEG signals to the $[0.5 : 8]Hz$ subband, since in [7] we have shown that this particular frequency range guarantees better recognition performance for VEP-based EEG biometrics, and then down-sample the signals from $256Hz$ to $128Hz$. Finally, the EEG signals are normalized using a “z-score” transformation, which generates zero-mean data with unitary variance. Eventually, each subjects' signals is detrended by individually subtracting its best-fit line, which allows us to concentrate on the data fluctuations of the estimated trend. Figure 2 shows all the preprocessing steps in sequential manner.

C. Convolutional Neural Network

A convolutional neural network (CNN) is a multilayer perceptron (MLP) network with a special topology containing more than one hidden layer [6]. CNN is primarily used for object recognition in image processing, handwritten character and speech recognition, as it automatically extracts discriminative features inside its layers from the raw input information, without any specific normalization. This kind of model is advantageous for input data with an inner structure like for instance images, and where invariant features have to be discovered. Such capability may be useful for dealing with EEG signals, which substantially vary over time and individual in their raw form, being therefore local-kernel-based architectures typically inefficient for classification purposes, since in most of the occasions it is not easy to determine the type of features that are supposed to be extracted. On the other hand, a CNN-based classifier can be an interesting approach for EEG clustering, as it may turn out to be more appropriate to let the network extract the most discriminant features by constructing high-level features through its back-propagation steps.

It has to be mentioned that CNN has already been proposed for EEG-based biometric recognition in [15], where however only brain signals acquired in resting states conditions have been evaluated. More importantly, the tests there performed have been based on EEG data acquired during a single session from only 10 users, being the resulting reliability questionable, as already commented in Section II for VEP-based approaches.

D. Network Topology

Our CNN network topology is shown in Figure 3. This network has 4 convolutional layers, 2 max-pooling, 1 ReLU, and a softmaxloss layer. The detailed network topology is described as follows:

- L_0 : The input layer with an input data size of $[17 \times 77]$, where 17 is the number of EEG acquisition channels and 77 represents 600ms signal after the display of a geometric shape image at a rate of $128Hz$, as described in Section III-A.
- L_1M_1 : First hidden layer, composed of 77 convolutional filter of size $[5 \times 5 \times 1]$ and a max-pooling (MP) layer of size $[2 \times 2]$. This layer transforms the input data into a size of $CL_1M_1 = [6 \times 36 \times 77]$ after convolving and down-sampling.
- L_2M_2 : Second hidden layer, which is composed of 320 conv filter of size $[5 \times 5 \times 77]$ and a max-pooling layer of size $[2 \times 2]$. This layer transforms the first hidden layer's output into a size of $CL_2M_2 = [1 \times 16 \times 320]$ high level features.
- $L_3M_2R_1$: This hidden layer is composed of 1024 convolutional filter of size $[1 \times 16 \times 320]$ and an Rectified Linear Unit (ReLU) layer, whose purpose is to introduce non-linearity into the system. This layer changes the previous layer's activation map into a $CL_3M_2R_1 = [1 \times 1 \times 1024]$ feature map.
- $L_4M_2R_1$: The output layer or the fully connected layer is produced by convolving the previous layer's activation map using 40 convolutional filter of size $[1 \times 1 \times 1024]$. This layer has only one map of 40 neurons, which represents the 40 classes/subjects. This layer is fully connected with $L_3M_2R_1$. Softmaxloss function is used here as a loss function for back-propagation.

IV. EMPLOYED EEG-BASED BIOMETRIC SYSTEM

Once EEG data are acquired and preprocessed, the corresponding templates are generated as described in Section IV-A. The performed training and identification network phases are described in Section IV-B and IV-C, respectively.

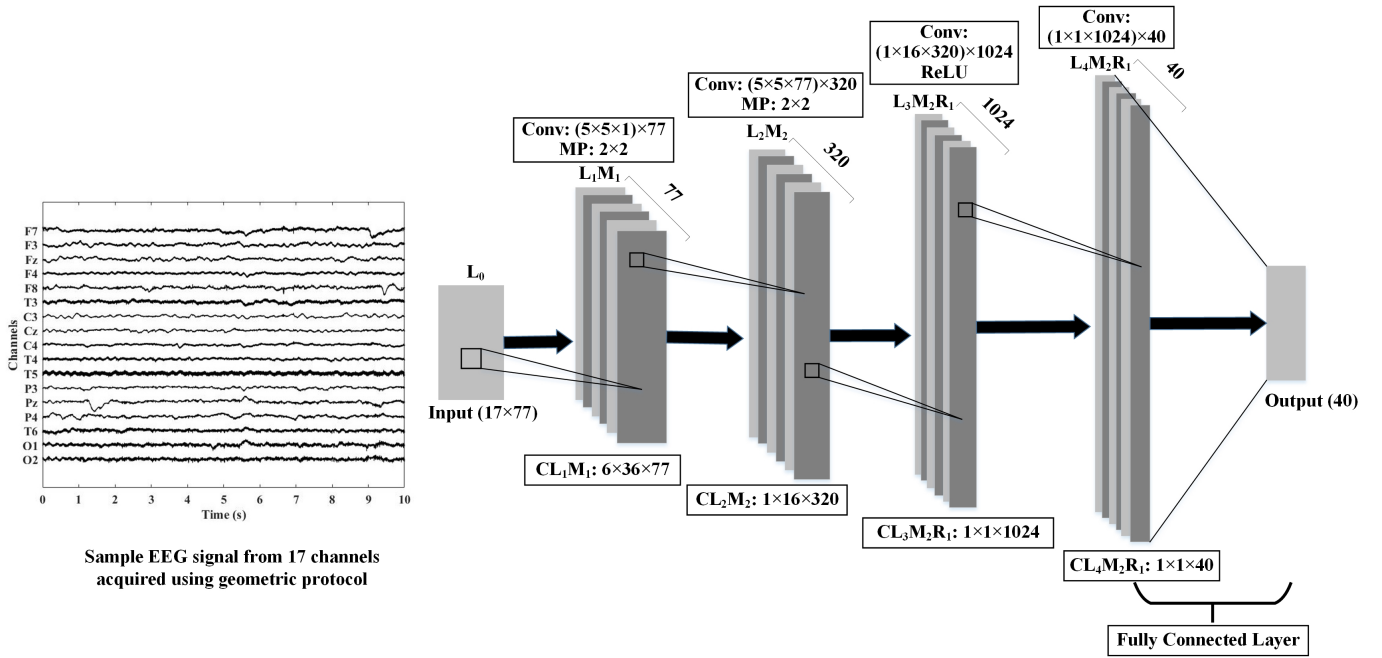


Fig. 3. Employed CNN Architecture with sample EEG signal acquired from 17 channels

A. Template Generation

Templates are generated from the acquired EEG data through a signal averaging process, as VEPs are usually significantly low in amplitude with respect to the overall behavior of EEG fluctuations, and therefore need to be evaluated over multiple repetitions in order to be extracted from the background. Since both target and non-target stimuli are time-locked with the originating events, we collect 60 responses of each stimuli, lasting for $T = 600ms$ from the outset of the associated event. For a particular user, the generated template is a collection of 17 time-dependent potentials, registered from $M = 17$ EEG channels in correspondence to either target or non-target stimuli [7]. Given an M -channel EEG signal collected through the proposed protocol, mean behaviors across $R = 50$ consecutive responses (out of 60 available for target events, and $7 \cdot 60$ for non-target events) 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 subject from $S1$ (training) and $S2$ (testing) session's data.

B. CNN Training

After the execution of the preprocessing steps, the EEG signals are passed through the newly designed CNN network, and inside the first hidden layer a set of very low level features are extracted. In the subsequent convolutional layers, the network gradually builds up over these low-level features, in order to create high-level features for fully connected layer.

In more detail, templates generated from session $S1$ are used for CNN training, meaning that, for every subject, we have an input data of size 17×77 , passed through the newly designed CNN network as described in Section III-D. For

CNN network designing and training we use the MatConvNet-1.0-beta16 tool [16]. The actual length of our training dataset is $17 \times 77 \times 2000$, where 40 subject's $T = 50$ templates generates 2000 samples. 90% of each subject's data were used for training purposes and rest 10% for validation. The learning rate of the CNN network is set at 0.001 with a batch size of 5 samples, so that the loss can be minimized with higher precision with the execution of every epochs. As for the number of epochs to be considered, higher numbers usually allow the network to get well trained so that the weights of different layers are updated with precision. For our experiment, we have investigated 20, 50 and 100 epochs and found that 50 epochs are enough to achieve higher accuracy.

C. Identification

In the identification stage, the testing templates are generated as described in Section IV-A from the $S2$ session. The testing dataset size is the same as the training dataset, where the only significant difference is that the training and testing datasets are from different acquisition sessions. For each testing sample of size 17×77 , the trained CNN network returns probability values corresponding to all the 40 classes/subjects. The maximum probability value identifies the subject with which the testing sample is more similar.

V. RESULTS & DISCUSSION

Experiments are performed for both target and non-target events separately, in order to discover the most suitable scheme for biometric identification based on VEP. Specifically, the accuracy of our CNN-network-based biometric system is evaluated through rank-wise identification rates. The results obtained for non-target vs. non-target scheme is shown in Figure 4. As can be seen, a 98.8% accuracy is achieved

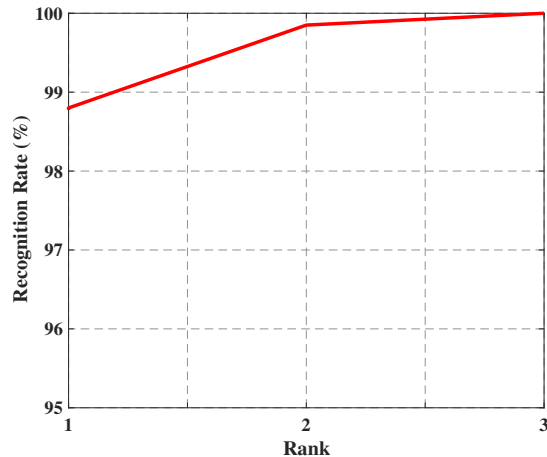


Fig. 4. “Geometric” protocol: Cumulative Match Curve (CMC) for M=17 Ch. combination of NonTarget vs. NonTarget Scheme.

for rank-1, and absolute accuracy is attained at rank-3. On the other hand, for the target vs. target scheme the rank-1 performance is at 80.65% accuracy, while 90% and 97% accuracies are respectively reached for rank-5 and rank-10. Figure 5 displays the identification performance for the target vs. target scheme.

Besides notably improving the recognition performance previously achieved in [7], where the VEP generated by the acquired EEG signals are compared through a simple cosine distance classifier, the above discussed results re-enforce our previously-stated claim, observing that the non-target vs. non-target scheme is able to achieve higher accuracy rate for biometric identification.

VI. CONCLUSIONS

In this paper we have proposed a visual-stimuli elicited EEG-based biometric identification mechanism, with a CNN is used to extract discriminative features and achieve a high degree of accuracy. Matching schemes based on non-target vs. non-target events perform much better than those relying on target vs. target events. The proposed CNN-network-based identification system, with its high recognition accuracy, allows us to take EEG biometrics into serious consideration for further investigation. In addition to its high level of security and confidentiality, EEG-based biometric system might be useful for physically disabled people, who are unable to use the conventional biometric systems like fingerprints, retinal scans etc. In future, EEG biometrics can have a far-reaching applications to different fields such as law enforcement, defense systems and others.

REFERENCES

- [1] H. H. Stassen, “Computerized recognition of persons by EEG spectral patterns,” vol. 49, 1980.
- [2] P. Campisi and D. La Rocca, “Brain waves for automatic biometric-based user recognition,” *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 5, pp. 782–800, 2014.

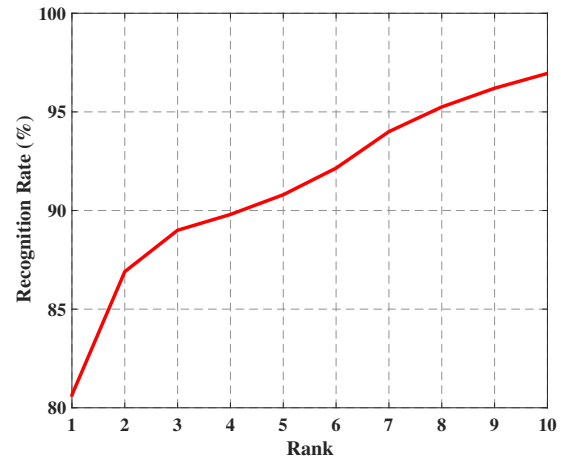


Fig. 5. “Geometric” protocol: Cumulative Match Curve (CMC) for M=17 Ch. combination of Target vs. Target Scheme.

- [3] A. Subasi and E. Erelebi, “Classification of EEG signals using neural network and logistic regression,” *Computer Methods and Programs in Biomedicine*, vol. 78, no. 2, pp. 87 – 99, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0169260705000246>
- [4] S. Solhjoo, A. M. Nasrabadi, and M. R. H. Golpayegani, “Classification of chaotic signals using hmm classifiers: EEG-based mental task classification,” in *2005 13th European Signal Processing Conference*, Sept 2005, pp. 1–4.
- [5] I. Guler and E. D. Ubeyli, “Multiclass support vector machines for EEG-signals classification,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 2, pp. 117–126, March 2007.
- [6] 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.
- [7] R. Das, E. Maiorana, and P. Campisi, “EEG biometrics using visual stimuli: A longitudinal study,” *IEEE Signal Processing Letters*, vol. 23, no. 3, pp. 341–345, March 2016.
- [8] R. Palaniappan, “Method of identifying individuals using vep signals and neural network,” in *Science, Measurement and Technology, IEE Proceedings-*, vol. 151, no. 1. IET, 2004, pp. 16–20.
- [9] H. Touyama, “EEG-based personal identification,” in *Biomedical Engineering*. InTech, oct 2009. [Online]. Available: <http://dx.doi.org/10.5772/7858>
- [10] C. N. Gupta, R. Palaniappan, and R. Paramesran, “Exploiting the p300 paradigm for cognitive biometrics,” *International Journal of Cognitive Biometrics*, vol. 1, no. 1, pp. 26–38, 2012.
- [11] K. Das, S. Zhang, B. Giesbrecht, and M. P. Eckstein, “Using rapid visually evoked EEG activity for person identification,” in *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*. IEEE, 2009, pp. 2490–2493.
- [12] S.-K. Yeom, H.-I. Suk, and S.-W. Lee, “Person authentication from neural activity of face-specific visual self-representation,” *Pattern Recognition*, vol. 46, no. 4, pp. 1159–1169, 2013.
- [13] B. C. Armstrong, M. V. Ruiz-Blondet, N. Khalifian, K. J. Kurtz, Z. Jin, and S. Laszlo, “Brainprint: Assessing the uniqueness, collectability, and permanence of a novel method for erp biometrics,” *Neurocomputing*, vol. 166, pp. 59–67, 2015.
- [14] J. Malmivuo and R. Plonsey, *Bioelectromagnetism : principles and applications of bioelectric and biomagnetic fields*. New York: Oxford University Press, 1995. [Online]. Available: <http://opac.inria.fr/record=b1084865>
- [15] 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.
- [16] A. Vedaldi and K. Lenc, “Matconvnet – convolutional neural networks for matlab,” in *Proceeding of the ACM Int. Conf. on Multimedia*, 2015.