

Research Article

Factor Analysis for Finding Invariant Neural Descriptors of Human Emotions

Vitor Pereira,¹ Filipe Tavares,¹ Petya Mihaylova,²
Valeri Mladenov ,² and Petia Georgieva ¹

¹Department of Electronics Telecommunications and Informatics/IEETA, University of Aveiro, Aveiro, Portugal

²Technical University of Sofia, Sofia, Bulgaria

Correspondence should be addressed to Petia Georgieva; petia@ua.pt

Received 2 August 2017; Accepted 16 January 2018; Published 19 March 2018

Academic Editor: Vittorio Loreto

Copyright © 2018 Vitor Pereira et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A major challenge in decoding human emotions from electroencephalogram (EEG) data is finding representations that are invariant to inter- and intrasubject differences. Most of the previous studies are focused in building an individual discrimination model for every subject (subject dependent model). Building subject-independent models is a harder problem due to the high data variability between different subjects and different experiments with the same subject. This paper explores, for the first time, the Factor Analysis as an efficient technique to extract temporal and spatial EEG features suitable to build brain-computer interface for decoding human emotions across various subjects. Our findings show that early waves (temporal window of 200–400 ms after the stimulus onset) carry more information about the valence of the emotion. Also, spatial location of features, with a stronger impact on the emotional valence, occurs in the parietal and occipital regions of the brain. All discrimination models (NN, SVM, kNN, and RF) demonstrate better discrimination rate of the positive valence. These results match closely experimental psychology hypothesis that, during early periods after the stimulus presentation, the brain response—to images with highly positive valence—is stronger.

1. Introduction

Electroencephalography (EEG) is an efficient noninvasive technique to analyze brain activity by measuring the electrical activity on the surface of the subject's scalp. EEG is particularly useful for development of brain-computer interfaces (BCIs). BCI refers to the direct communication pathway between the brain and an external device, such as a computer. In this paper, we focus on a subarea of BCI known as affective computing, which studies the neural mechanisms of emotions. The goal of the present work is to reliably discriminate between human emotions with positive or negative valence based on extracted EEG neural signatures. The valence, when used in psychology, refers to the “positiveness” or “negativeness” perceived by a person when exposed to some stimulus or events.

Working with EEG data brings several challenges. Brain waves recorded in the EEG have a very low signal-to-noise ratio and the noise comes from a variety of sources. For instance, the sensitive recording equipment can easily

pick up electrical line noise from the surroundings. Other unwanted electrical noise can come from muscle activity, eye movements, or blinks. The EEG lacks spatial resolution; however it has a good (millisecond) time resolution to record both slowly and rapidly changing dynamics of the brain activity. Therefore, in order to identify the brain activity of interest, the relevant content of the EEG signal needs to be separated from the noise and the background processes.

EEG offers many advantages for construction of a BCI system but also several disadvantages. Most importantly, EEG is a noninvasive method for measuring brain activity. This removes the need for costly and risky surgical procedures, such as electrophysiology, in which intracortical devices such as needles or tubes may be inserted directly into the brain material, or electrocorticography, in which an array of electrodes is implanted under the skull. Both systems risk permanent and life threatening damage to a patient's brain and require costly surgical expertise to carry them out safely. Also useful for designing a BCI, EEG does not require the

patient to be stationary like other noninvasive imaging systems such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG), both of which can only be carried out by large scale and expensive equipment. EEG has the ability to produce high time resolution data, which is a necessity for near-real-time systems.

Research on methods of extracting emotions from EEG patterns has intensified over the last decade. A comprehensive list of EEG-based emotion recognition systems is provided in [1]. Effective features that correlate EEG patterns and emotions are discovered both in time and in frequency domains. In the frequency domain, the EEG energy and power, the Power Spectral Density (PSD), and the Spectral Power Asymmetry (ASM) of certain bands, such as beta (16–32 Hz) and gamma (32–64 Hz), are most frequently selected. In the time domain, the Fractal Dimension (FD), Higher Order Crossings (HOC), and sample entropy have been successfully utilized [2]. More recent strategies propose Deep Learning networks to extract high-level representations of the raw EEG signals [3, 4]. To alleviate overfitting problem, principal component analysis (PCA) is applied to extract the most important components of the initial feature set. Hybrid approaches using multidimensional information, such as EEG and facial expression [5] and EEG combined with Biometric and Eye Tracking Technologies [6], proved to outperform the single modality (only EEG signals) emotion classification.

The refereed studies usually build an individual discrimination model for every subject (subject dependent model). Building subject-independent models is a harder problem due to the high data variability between different subjects and different experiments with the same subject. The extracted features improve the emotion discrimination performance indeed; however their computational nature (e.g., entropy, FD, and wavelet transform) turn to be more difficult to be interpreted by nonexperts in machine learning and data mining.

The contribution of this paper is twofold. First, we propose a new spatial-temporal feature set with clear physical ground, that is, local amplitudes and their latencies (time of occurrence) over selected EEG electrodes. Second, we propose, for the first time in affective computing, the Factor Analysis (FA) as an alternative method to extract the most important components of the feature set. In contrast to the widely used PCA, FA preserves the relation between the factors and the feature set and therefore the conclusions are more easily interpreted by psychologists and neurophysiologists. The combination of the proposed features and FA resulted into finding the most representative neural descriptors of the emotion valence, invariant with respect to trials and subjects.

Many of the previous studies are based on the publicly available dataset for emotion analysis DEAP, where EEG and peripheral physiological signals of 32 subjects are simultaneously recorded as they watched music videos [7]. In the present work, data generated during experiments in the University of Aveiro are used. Time-locked EEG signals, a.k.a. Event-Related Potentials (ERPs), were registered during periods when the subjects have been exposed to high arousal images [8]. Though the present results were obtained with

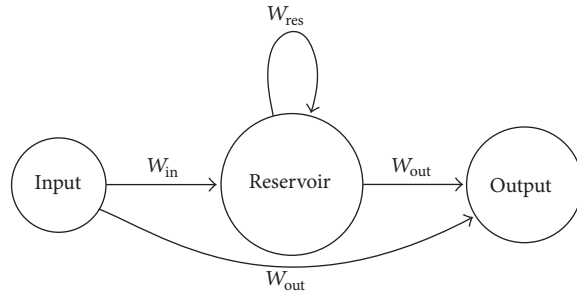


FIGURE 1: Echo State Network (ESN) architecture. Image by C. Areshenkoff. Obtained from <http://areshenk-research-notes.com/echo-state-networks-in-r/>.

different data, our findings support previous neurophysiological studies regarding the correlation between early and late EEG waves (P1, N1, P2, and N2 components) and their respective latencies with human emotional states [9–11]. This is also a valuable contribution to the efforts of understanding the neural correlations of the emotions.

This paper is organized in the following way. In Section 2 an overview of state-of-the-art feature extraction techniques is provided. In Section 3 the experimental dataset is described. In Section 4 the selected discrimination models are detailed. The results of decoding human emotions based on the extracted neural descriptors are discussed in Section 5. The conclusions are summarized in Section 6.

2. Invariant Feature Extraction

In this section some of the most successfully applied feature extraction and dimensionality reduction techniques in multisubject data framework are reviewed.

2.1. Echo State Networks for Feature Extraction. Bozhkov et al. [12] introduced a feature extraction method using Echo State Networks (ESN). ESN is a particular type of Recurrent Neural Networks, also known as a *reservoir computing*, Figure 1. The connections in the reservoir are initialized randomly and remain unaltered during the ESN’s training. The input layer is connected to the reservoir and to the output layer and its weights are also randomly initialized and untrained. Only the weights connecting the output of the reservoir to the output layer are trained [13]. The computational advantage of the ESN due to lower number of training weights limits its capacity to extract the underlying data characteristics.

A modification of the ESN is proposed in [14] to fix this problem, where the reservoir weights are iteratively updated applying intrinsic plasticity (IP) adaptation rule. IP computes the equilibrium states of the reservoir neurons and this adaptation step greatly improves the low dimensional projections of the equilibrium states at the output layer.

2.2. Interval Feature Extraction. Kuncheva and Rodríguez [15] introduced interval feature extraction (IFE) which consists of extracting features derived from the signal over time segments of various lengths.

If x_i is the value of the signal at time i where $i = 1, \dots, T$ spans the ERP episode then nested time intervals are formed

for each i using the remaining samples x_i, \dots, x_T . The interval length is varied, and, for example, a sequence of such intervals can be $\{x_i, x_{i+1}\}, \{x_i, x_{i+1}, x_{i+2}, x_{i+3}\}, \dots, \{x_{i+1}, x_{i+2}\}$, ending with $\{x_{T-1}, x_T\}$. The following quantities are extracted from each interval:

- (i) Average amplitude: μ
- (ii) Standard deviation: σ
- (iii) Covariance:

$$\gamma = \frac{1}{l} \sum_{k=i_b}^{i_e} kx_k - \mu \frac{i_b + i_e}{2}, \quad (1)$$

where i_b and i_e are the starting and ending time indexes and the length of the interval is $l = i_e - i_b + 1$.

2.3. Independent Component Analysis. Independent component analysis (ICA) is typically used to “clean up” EEG data by removing noisy electrodes, blinks, and other artefacts. ICA produces a matrix s of maximally independent components when multiplying an unmixing linear static transformation matrix W to the original data x .

$$s = Wx. \quad (2)$$

Stewart et al. [16] combine ICA with support vector machine (SVM) to classify whether the EEG readings indicate the presence of visual object stimuli. Mueller et al. [17] also use ICA and SVM to discriminate ADHD versus healthy adults through EEG data captured during a Go/NoGo task. In both studies [16, 17] the matrix W is found applying Infomax algorithm [18]. Infomax maximizes the joint entropy of nonlinearly transformed ensembles of zero-mean input vectors.

Solving the previous equation for x ,

$$x = W^{-1}s. \quad (3)$$

Further,

$$x = \sum x_i = \sum W_i^{-1}s_i, \quad (4)$$

where s_i is the i th independent component and W_i^{-1} is its respective cofactor. s_i can be obtained as

$$s_i = W_{zi}x. \quad (5)$$

W_{zi} is a sparse matrix with all entries equal to zero except the i th row. Finally,

$$x_i = W^{-1}W_{zi}x. \quad (6)$$

$W^{-1}W_{zi}$ is the filter that projects the original data into the space of the independent components. This filter was used to decompose individual ERP waves into individual components.

2.4. Graded Probabilistic Clustering. Masulli et al. [19] proposed a fuzzy Graded Probabilistic Clustering (GPC-II), where each data point x_i can belong to more than one cluster. The centroids y_j of the clusters are given as the mean of all points weighed by their degree (membership) of belonging to the cluster u_{ij} :

$$y_j = \frac{\sum_i u_{ij}x_i}{\sum_i u_{ij}}. \quad (7)$$

GPC-II was applied in experiments involving Go/NoGo task where the participants were shown pictures of faces and asked to press a button when the face shows “expressed neutral emotion.” The goal was to find clusters in the recorded multisubject EEG data. GPC-II runs with a high number of clusters and at each iteration a cluster that covers only a single trial (a “singleton”) is removed. Neighbor clusters are merged based on the Jaccard index as a measure of closeness of pair of clusters. Two clusters are merged if their index surpasses a threshold Jaccard index value. The new centroid is computed as the average of the centroids of the previous clusters weighed by the membership values of their elements. This procedure is applied iteratively until no more neighboring clusters can be merged.

2.5. Optimal Feature Selection. Optimal feature selection (OFS) is a deterministic greedy algorithm that takes the locally optimal choice at each stage [20]. OFS can be implemented in two ways: forward selection or backward elimination.

Forward Selection. Build data models based on individual features. Choose the best $k1$ features. Make all two by two combinations of the $k1$ selected features and build new data models. Choose the next group of best $k2$ features and make all three by three combinations of them. Repeat the process until there is no more improvement in the model performance. The choice of subsequent feature subset combinations may vary in order to reflect the experience gained during the extraction process and the application in hand.

Backward Elimination. The model is trained on the complete set of features and weights are assigned to each of them. The features with the smallest weights are then pruned from the set. This process is repeated until the model performance deteriorates below a certain threshold.

2.6. Principal Component Analysis. Generally, Principal Component Analysis (PCA) refers to the statistical process used to emphasize variation for which principal data components are computed and bring out strong patterns in the dataset. It is often used to make data easy to explore and to visualize everything in a single 2D graph (usually called a biplot).

Consider that $X = [X_1, \dots, X_p]$ are the p random variables under study. PCA will consist of obtaining p linear combinations of these p variables as follows:

$$\begin{aligned} Y_1 &= a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \\ Y_2 &= a_{12}X_1 + a_{22}X_2 + \dots + a_{p2}X_p \\ &\dots \\ Y_p &= a_{1p}X_1 + a_{2p}X_2 + \dots + a_{pp}X_p \end{aligned} \quad (8)$$

in such a way that $\text{Var}(Y_1) \geq \text{Var}(Y_2) \geq \dots \geq \text{Var}(Y_p)$ and all pairs (Y_i, Y_j) , with $i \neq j$, are uncorrelated variables.

In order to perform a PCA over the dataset, it is necessary to obtain the eigenvalues (and the eigenvectors) of the covariance matrix.

2.7. Factor Analysis. Factor Analysis (FA) is an explorative method similar to PCA. Much like the cluster analysis of grouping similar cases, the FA groups similar variables into the same factor [21]. This process is usually referred to as identifying latent variables. Due to its explorative nature, it does not distinguish between independent and dependent variables; it only uses the data correlation matrix.

In other words, FA reduces the information in a model by decreasing the dimension of the dataset. This procedure can have multiple purposes. It can be used to simplify the dataset, for example, reducing the number of variables in predictive regression models. If Factor Analysis is used for these purposes, normally the factors are rotated after its extraction. FA considers several different rotation methods that ensure that the factors are orthogonal. Therefore, the correlation coefficient between two factors is exactly zero. For example, it totally eliminates problems of multicollinearity in regression analysis.

The most commonly used rotation method is *Varimax*, which is the one that will be used in the following section. *Varimax* consists in an orthogonal rotation method (which produces independent factors) that minimizes the number of variables that have high loadings on each factor. The objective of this method is to simplify the interpretation of each one of the factors.

Steps in Exploratory Factor Analysis

- (i) Collect and explore data: choose relevant variables.
- (ii) Extract initial factors (via principal components).
- (iii) Choose number of factors to retain.
- (iv) Choose estimation method and estimate model.
- (v) Rotate and interpret the results.

Factor Analysis Mathematical Model

$$\mathbf{X}_{(p \times 1)} = \mathbf{L}_{(p \times m)} \times \mathbf{F}_{(m \times 1)} + \boldsymbol{\varepsilon}_{(p \times 1)}, \quad (9)$$

where each matrix represents the following:

$\mathbf{X}_{(p \times 1)}$: multivariate random vector with p variables

$\mathbf{F}' = [F_1 \ F_2 \ \dots \ F_m]$: factors vector with $m < p$ factors common to all p random variables.

$\mathbf{L}_{(p \times m)}$: matrix of loadings of the factors; the coefficient l_{ij} represents the weight of the i th variable to j th factor

$\boldsymbol{\varepsilon}' = [\varepsilon_1 \ \dots \ \varepsilon_p]$: vector of errors.

To sum up, the Factorial Analysis Model considers that the variables could be grouped by their correlations. It is expected that when a high correlation exists between two variables they will be related to the same factor (in the sense that both will have high loadings on that factor).

3. Data Description and Preprocessing

In this section we describe the data used to demonstrate the search for common neural signatures feasible to decode human emotions of various subjects.

ERPs were collected while 26 female volunteers were exposed to 24 high arousal images with positive or negative content from the International Affective Picture System [22]. Each image was shown 3 times in pseudorandom order and each trial lasted 3500 ms. 12 features (amp1, amp2, amp3, amp4, amp5, amp6, latency1, latency2, latency3, latency4, latency5, and latency6) were extracted from 21 channels of recorded ERPs, positioned according to the 10–12 system, and 2 EOG channels were sampled at 1000 Hz. The maximum and minimum values of the ensemble average signals, computed and filtered using a zero-phase filtering scheme, were detected and the features correspond to the time and amplitude characteristics of the first three minimums and maximums occurring after $T = 0$ [12]. The starting data matrix is composed of 252 ($12 * 21$) columns of features and 52 ($2 * 26$) rows of examples. The examples correspond to the averaged class associated data of each participant. The class of the examples is defined by the positive (1) or negative (0) content of the presented image. Note that the number of features is much higher than the number of examples and therefore feature reduction is strongly recommended.

The recording equipment NeuroScan provides initial filtering, eye-movement correction, baseline compensation, and division of the signals in epochs. Prior to feature extraction the recorded EEG signals were further filtered with 4th-order Butterworth filter with passband [0.5–15] Hz [12]. Data were also normalized as follows:

$$z = \frac{x - \mu}{\sigma}, \quad (10)$$

where x is the original data point and μ and σ are the mean and the standard deviation of the data distribution.

4. Discrimination Models

Typical discrimination models were selected—Neural Network (NN), Random Forest (RF), Support Vector Machine (SVM), and k -Nearest Neighbors (kNN). This choice reflects the focus of the present study in finding suitable neural descriptors, less dependent on the classifier characteristics. Data were split into training and validation subset (76%) and

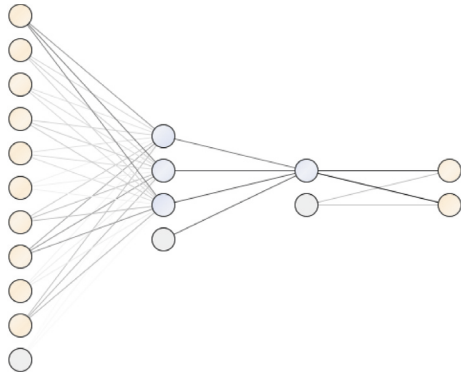


FIGURE 2: Neural network model.

testing subset (24%). k -fold cross validation (CV) was used with $k = 26$. At each CV iteration, 25 subjects were used for training and a single subject (her positive and negative valence readings) was used for validation. First, emotion detection based on all 252 features was performed and then the feature selection techniques presented in Section 2 were applied prior to the emotion detection. The optimal structure and hyperparameters of the classifiers were chosen after a grid search over a sequence of k -fold CV as described below.

Neural Network (NN)

- (i) Fully connected NN with two hidden layers: the first one with 3 neurons and the second layer with one neuron (Figure 2).
- (ii) The learning rate was optimized in the range of [0.01 to 1.0], with the optimal value equal to 0.3.
- (iii) The momentum was optimized in the range of [0.0 to 1.0], with the optimal value equal to 0.2.

SVM

- (i) SVM with dot-type of kernel ($k(x, y) = x * y$) was used.
- (ii) C was optimized in the range of [0.1 to 10], with the optimal value equal to 1.

The performance of the SVM was improved applying the adaptive boosting meta-algorithm (*AdaBoost*). *AdaBoost* fits additional copies of the classifier on the dataset with adjusted weights on the incorrectly classified instances so as to focus the classifier on the most difficult to classify instances [23].

K -Nearest Neighbors (k NN)

- (i) k was optimized in the range of [1 to 10], with the optimal value equal to 5.
- (ii) Distance measures for computing the k -Nearest Neighbors. Two distance measures were comparatively studied:

- (1) Chebyshev distance between two points p and q , with coordinates p_i and q_i , is

$$D_{\text{Chebyshev}}(p, q) := \max_i (\text{abs}(p_i - q_i)). \quad (11)$$

- (2) Cosine similarity is a measure of similarity between two nonzero vectors expressed by the cosine of the angle between them [24]. The results in the next section are obtained applying k NN with cosine similarity as this distance measure proves to provide better discrimination.

Random Forest. The Random Forest (RF) consists of multiple decision trees. We run RF with {5, 10, 15, 20} trees and maximum tree depth of {10, 20, 50, 80}. The optimal RF structure was found with 10 trees and maximum depth of 20.

5. Results

In this section the results obtained with the statistical software RapidMiner (RM) are presented, extended with RM Feature Selection [25]. The experiments run on a computer running Windows 10 with an Intel Core i7-5500u CPU and 8 GB of RAM memory.

The performance of all discrimination models on test data is summarized in Table 1. The emotion valence recognition across subjects is far too low when all features are provided as inputs to the classifiers (the second column in Table 1). This is somehow an expected result taking into account the high dimension of the feature space compared to the number of examples. Running the classification task after a preprocessing step with any of the feature selection techniques presented in Section 2 improves the performance compared to a direct brain state discrimination. However, the Factor Analysis (FA) is revealed to be the winning approach in extracting invariant features. Among the four classifiers, SVM and NN present the highest recognition rate.

Table 2 presents the Factor Analysis performed using packages in *software R* library, considering four factors and the standard *Varimax* rotation [26]. FA establishes the general relationship between variables. The first four factors explain about 70% of the total variability of the 252 initial variables. The table shows the strongest relation between factors and original data variables and can be interpreted as follows:

- (i) *Temporal features (amplitude, latency)*: early waves (Lmin1, Lmax2, i.e., temporal window of 200–400 ms after the stimulus onset) have higher influence on data variability. This matches experimental psychology hypothesis related to temporal dynamics of emotions. According to [27], early waves carry more information about the valence than the arousal of the emotion. Late waves are less discriminative with respect to the emotion valence.
- (ii) *Spatial features (channels)*: spatial location of features, with stronger impact on the emotional valence, occur in the parietal (channels P3, Pz, P4, and P7) and occipital (channels O1, Oz, and O2) regions of the brain.

TABLE 1: Test Accuracy of 4 classifiers (NN, SVM, RF, and kNN) with feature selection methods: ESN: Echo State Network; IFE: Interval Feature Extraction; ICA: Independent Component Analysis; GPC: Graded Probabilistic Clustering; OFS: Optimal Feature Selection; Factor Analysis (FA).

	-	ESN	IFE	ICA	GPC	OFS	FA
NN	75,0	85,9	78,1	81,8	83,5	80,3	90.4
SVM	58,3	86,7	77,1	72,9	64,2	73,6	90.4
RF	50,0	79,7	71,3	69,4	62,8	65,2	84.6
kNN	56,7	83,3	72,2	72,6	76,8	79,5	88.5

TABLE 2: Factor Analysis loadings for *Varimax* rotated matrix of four-factor model explaining about 70% of the total dataset variance (presented in bold if absolute values are greater than 0.7).

	F1	F2	F3	F4
P3 (Amin2)	0.91	-0.13	0.07	0.02
Pz (Lmin1)	0.78	0.10	-0.44	0.01
P4 (Lmax2)	0.75	0.04	0.44	0.23
P7 (Amin1)	-0.73	0.01	-0.01	-0.25
O1 (Amin2)	-0.14	0.88	0.03	0.07
Oz (Lmax2)	-0.02	0.88	-0.11	0.09
P8 (Amax2)	0.35	0.46	-0.03	-0.25
C3 (Amin2)	-0.23	-0.11	0.78	-0.13
O2 (Lmin1)	-0.34	-0.02	-0.72	-0.13
P3 (Amin2)	0.13	0.00	-0.27	0.80
P7 (Amin1)	0.15	0.06	0.27	0.76
Variance (%)	26	16	15	13

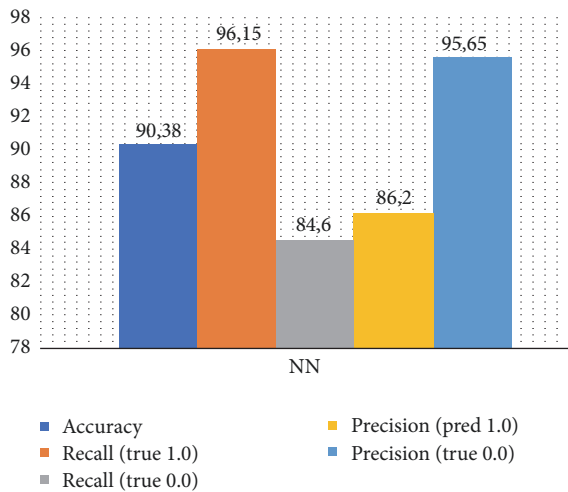


FIGURE 3: Neural Network (NN) performance measures with Factor Analysis (FA).

The discrimination rates in the last column of Table 1 were obtained with models trained with the four factors. More detailed analysis of the classifiers is depicted in Figures 3–6. The figures resemble the confusion matrix performance metrics. Note that the recall metric (the fraction of correctly classified positive examples) of all classifiers has the highest rate. This result corresponds to better discrimination rate of the positive valence, which is also in accordance with

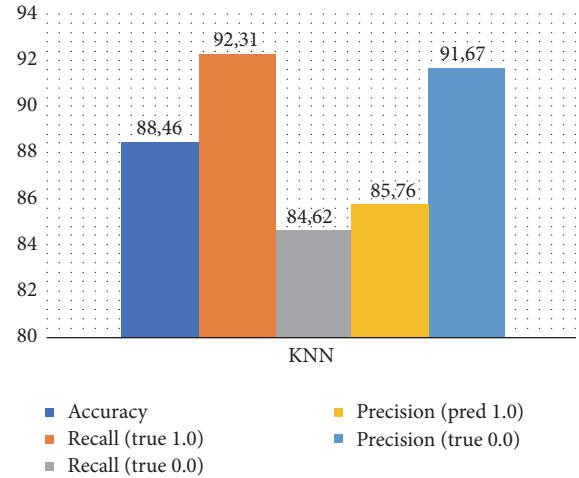


FIGURE 4: *k*-Nearest Neighbor (kNN) performance measures with Factor Analysis (FA).

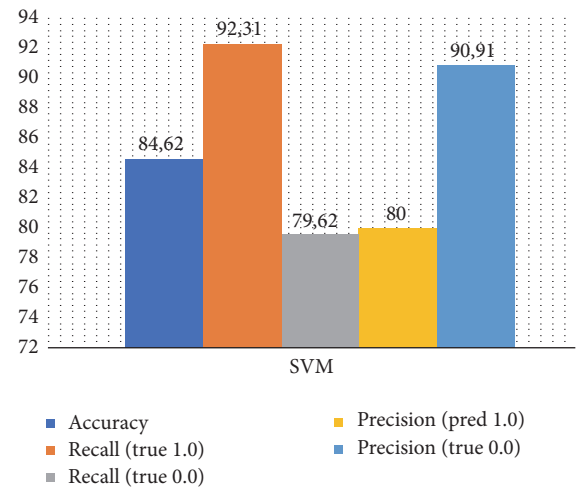


FIGURE 5: Random Forest (RF) performance measures with Factor Analysis (FA).

hypothesis of biological psychology [28] that during early periods after the stimulus presentation the brain response, to images with highly positive valence, is stronger.

6. Conclusion

The goal of this study was to identify common neural signatures based on which the positive and negative valence

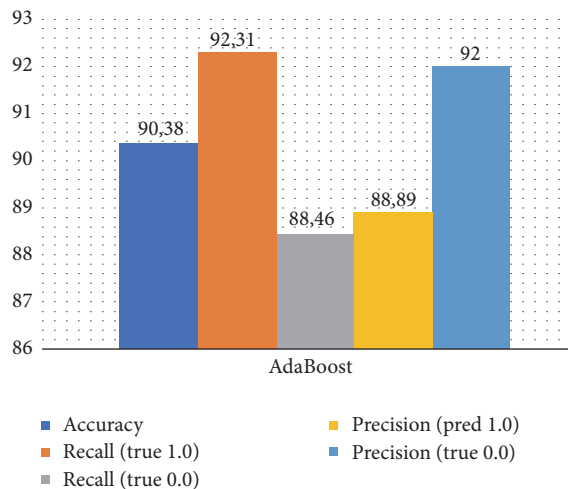


FIGURE 6: SVM performance measures with Factor Analysis (FA).

of human emotions across multiple subjects can be reliably discriminated. The brain activity is registered by Event-Related Potentials (ERPs). We explored the feasibility of training cross-subject discrimination models to make predictions based extracted invariant neural descriptors, hidden in the ERPs. The core of the present study is the way the features are selected. The combination of a small number of time domain (ERP amplitudes and latencies) and spatial (selected channels) features has the potential to reduce the intersubject variability and improve the learning of representative models valid across multiple subjects.

Based on the results of this study we believe that the Factor Analysis (FA) of ERPs is a promising approach to extract statistical underlying correlations of the brain activity among subjects and therefore decode human emotional states. Nevertheless, before making stronger conclusions on the capacity of the FA to decode emotions, further research is required to answer other questions such as discrimination of more than two emotions. In fact this is a relevant question for all reported works on affective neuroscience [1]. The discrimination is usually limited to two, three, or maximum four valence-arousal emotional classes. An interesting problem is also the human personality clustering based on EEG, for example, distinguishing between high versus low neurotic type of personality. Also, the number of participants in the experiments is important for revealing stable cross-subject features. In the reviewed literature the average number of participants is about 10–15, and the maximum is 32 (DEAP). We need higher dimensional datasets to compare different techniques in order to further progress the affective computing.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research work is funded by the Sofia Tex Park through R&D&I Consortium and Artificial Intelligence and CAD

Systems Lab, in the context of the project “Development and Implementation of Wireless Robot Control Concepts through EEG Signals from Brain Activity,” and the National Funds through the FCT (Foundation for Science and Technology), in the context of Project UID/CEC/00127/2013.

References

- [1] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, “Real-time EEG-based happiness detection system,” *The Scientific World Journal*, vol. 2013, Article ID 618649, pp. 1–12, 2013.
- [2] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, “Emotion recognition from EEG signals using multidimensional information in EMD domain,” *BioMed Research International*, vol. 2017, Article ID 8317357, pp. 1–9, 2017.
- [3] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, “EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation,” *The Scientific World Journal*, vol. 2014, Article ID 627892, pp. 1–10, 2014.
- [4] S. Ding, N. Zhang, X. Xu, L. Guo, and J. Zhang, “Deep extreme learning machine and its application in EEG classification,” *Mathematical Problems in Engineering*, vol. 2015, Article ID 129021, pp. 1–11, 2015.
- [5] Y. Huang, J. Yang, P. Liao, and J. Pan, “Fusion of facial expressions and EEG for multimodal emotion recognition,” *Computational Intelligence and Neuroscience*, vol. 2017, Article ID 2107451, pp. 1–8, 2017.
- [6] J. López-Gil, J. Virgili-Gomá, R. Gil et al., “Method for improving EEG based emotion recognition by combining it with synchronized biometric and eye tracking technologies in a non-invasive and low cost way,” *Frontiers in Computational Neuroscience*, vol. 10, article 85, 2016.
- [7] S. Koelstra, C. Mühl, M. Soleymani et al., “DEAP: a database for emotion analysis; using physiological signals,” *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [8] O. Georgieva, S. Milanov, P. Georgieva, I. M. Santos, A. T. Pereira, and C. F. Silva, “Learning to decode human emotions from event-related potentials,” *Neural Computing and Applications*, vol. 26, no. 3, pp. 573–580, 2015.
- [9] M.-K. Kim, M. Kim, E. Oh, and S.-P. Kim, “A review on the computational methods for emotional state estimation from the human EEG,” *Computational and Mathematical Methods in Medicine*, Article ID 573734, pp. 1–13, 2013.
- [10] J. K. Olofsson, S. Nordin, H. Sequeira, and J. Polich, “Affective picture processing: an integrative review of ERP findings,” *Biological Psychology*, vol. 77, no. 3, pp. 247–265, 2008.
- [11] J. K. Olofsson and J. Polich, “Affective visual event-related potentials: arousal, repetition, and time-on-task,” *Biological Psychology*, vol. 75, no. 1, pp. 101–108, 2007.
- [12] L. Bozhkov, P. Koprinkova-Hristova, and P. Georgieva, “Learning to decode human emotions with Echo State Networks,” *Neural Networks*, vol. 78, pp. 112–119, 2016.
- [13] H. Jaeger, “Echo state network,” *Scholarpedia*, vol. 2, no. 9, article 2330, 2007.
- [14] L. Bozhkov, P. Georgieva, I. Santos, A. Pereira, and C. Silva, “EEG-based subject independent affective computing models,” *Procedia Computer Science*, vol. 53, pp. 375–382, 2015.
- [15] L. I. Kuncheva and J. J. Rodríguez, “Interval feature extraction for classification of event-related potentials (ERP) in EEG data

- analysis,” *Progress in Artificial Intelligence*, vol. 2, no. 1, pp. 65–72, 2013.
- [16] A. X. Stewart, A. Nuthmann, and G. Sanguinetti, “Single-trial classification of EEG in a visual object task using ICA and machine learning,” *Journal of Neuroscience Methods*, vol. 228, pp. 1–14, 2014.
- [17] A. Mueller, G. Candrian, J. D. Kropotov, V. A. Ponomarev, and G. M. Baschera, “Classification of ADHD patients on the basis of independent ERP components using a machine learning system,” *Nonlinear Biomedical Physics*, vol. 4, no. 1, article S1, 2010.
- [18] S. Makeig, A. J. Bell, T. P. Jung, and T. J. Sejnowski, “Independent component analysis of electroencephalographic data,” *Advances in Neural Information Processing Systems*, vol. 8, no. 8, pp. 145–151, 1996.
- [19] P. Masulli, F. Masulli, S. Rovetta, A. Lintas, and A. E. P. Villa, “Unsupervised analysis of event-related potentials (ERPs) during an emotional Go/NoGo task,” in *Fuzzy Logic and Soft Computing Applications*, vol. 10147 of *Lecture Notes in Computer Science*, pp. 151–161, 2017.
- [20] L. Bozhkov and P. Georgieva, “Brain neural data analysis with feature space defined by descriptive statistics,” in *Pattern Recognition and Image Analysis*, vol. 9117 of *Lecture Notes in Computer Science*, pp. 415–422, Springer International Publishing, New York, NY, USA, 2015.
- [21] L. Tucker and R. MacCallum, *Exploratory Factor Analysis*, University of Illinois and Ohio State University, 1997.
- [22] “International Affective Picture System data,” <http://csea.phhp.ufl.edu/media/iapsmessage.html>.
- [23] R. E. Schapire, “Explaining AdaBoost,” in *Empirical inference*, pp. 37–52, Springer, Berlin, Germany, 2013.
- [24] Cosine similarity. (n. d.), In Wikipedia, https://en.wikipedia.org/wiki/Cosine_similarity.
- [25] <https://sourceforge.net/projects/rm-featselext/>.
- [26] O. O. Omo-Irabor, S. B. Olobaniyi, K. Oduyemi, and J. Akunna, “Surface and groundwater water quality assessment using multivariate analytical methods: a case study of the Western Niger Delta, Nigeria,” *Physics and Chemistry of the Earth*, vol. 33, no. 8–13, pp. 666–673, 2008.
- [27] M. Stolarova, A. Keil, and S. Moratti, “Modulation of the C1 visual event-related component by conditioned stimuli: evidence for sensory plasticity in early affective perception,” *Cerebral Cortex*, vol. 16, no. 6, pp. 876–887, 2006.
- [28] B. N. Cuthbert, H. T. Schupp, M. M. Bradley, N. Birbaumer, and P. J. Lang, “Brain potentials in affective picture processing: covariation with autonomic arousal and affective report,” *Biological Psychology*, vol. 52, no. 2, pp. 95–111, 2000.




Hindawi

Submit your manuscripts at
www.hindawi.com

